

修士論文

低解像度熱画像を用いたパターン認識の研究

11801639 萩原智宣

指導教員 伊藤直史

2013 年 3 月

群馬大学工学部電気電子工学科

情報通信システム第 4 研究室

目次

第1章 序論	4
1.1 非接触ユーザ・インターフェース	4
1.2 熱赤外センサと熱赤外画像	4
1.3 本研究の目的	5
第2章 熱赤外画像のパターン認識手法の原理	6
2.1 熱赤外データの前処理	8
2.2 ジェスチャーの抽出	9
2.3 離散コサイン変換を用いたパターン認識	11
2.3.1 離散コサイン変換	11
2.3.2 認識システムの流れ	12
2.3.3 評価式	13
2.4 フーリエ記述子を用いたパターン認識	14
2.4.1 フーリエ記述子の概要	14
2.4.2 輪郭抽出	17
2.4.3 代表点の抽出	19
2.4.4 評価式	21
第3章 離散コサイン変換を用いたパターン認識シミュレーション	22
3.1 使用する低解像度熱赤外画像	22
3.2 回転に対する DCT 係数の変化	27
3.3 手の傾きに対する DCT 係数の変化	32
第4章 フーリエ記述子を用いた回転に対するパターン認識シミュレーション	35
4.1 使用する低解像度熱赤外画像	35
4.2 輪郭抽出と代表点の抽出シミュレーション.....	39
4.3 画像の回転と代表点数を変化させた場合の認識シミュレーション...	47
4.3.1 Z 型記述子パターン認識を用いたシミュレーション.....	48
4.3.2 P 型記述子パターン認識を用いたシミュレーション.....	56

第5章 結論と展望	64
5.1 研究結果のまとめ	64
5.2 今後の展望	65
謝辞	66
参考文献	67

第1章

序論

1.1 非接触ユーザ・インターフェース

昨今、リモコンやスイッチに直接触れず、ジェスチャーなどを用いて操作を行う非接触ユーザ・インターフェースの開発が各所で進められている。最も身近なものとして Microsoft がゲーム機 XBOX360 の周辺機器として発売している「Kinect」が挙げられる。これは、可視光センサと距離画像センサを搭載した非接触ユーザ・インターフェースである。従来のゲーム機の操作方法は、リモコンやコントローラを用いてそれらから信号を送ることによって行われていたが、「Kinect」ではユーザのジェスチャーを検出することで操作を行い、何も持たずに信号を送ることが可能となっている。非接触ユーザ・インターフェースはそういった新しいゲームのプレイ環境を提供すると共に、物に触れずに操作出来るということから、衛生的観点から見ても有効であると考えられ、医用や防犯など様々な場所での利用が考えられている。非接触ユーザ・インターフェースの開発によって将来的にはリモコンを持つこともなく、その場のジェスチャーだけでコンピュータや家電機器を操作できるような、まるで SF 映画の世界観の生活環境を再現できる様になるかもしれない。

1.2 熱赤外センサと熱赤外画像

熱赤外センサは、対象物の温度を距離のある場所から非接触で測定できるセンサであり、照明の明るさに依らず観測することができる。例えば、夜間のリビングでスイッチを探すことなく、手で合図を送るだけで電灯を灯すようなシステムを作ることを考えると、暗闇であるから可視光センサでは難しく、熱赤外センサを採用するのが適当である。また、人体部分の温度はある一定の範囲内であることが多く、比較的容易に人体部分を抽出することができるというメリットがある。これらのことから、熱赤外センサを人間のジェスチャーを信号とした非接触ユーザ・インターフェースのハードウェアとして採用できるのではないかと考える。

しかしながら、熱赤外は可視光に比べると光の持つエネルギーが小さい。そのため、高精細のセンサを用いて対象を観測しようとした場合、観測できる熱放射が小さく、SN比を低下させずに高解像度の温度情報を取得することは中々難しい。そのため、現在販売されている熱赤外センサは、可視光のカメラに比べて非常に高価なものが多く 2500 画素程度のセンサでも 10 万円を超えてしまうのが現状である。このことから、熱赤外

センサをジェスチャー認識に対応する汎用的な非接触ユーザ・インターフェースとして採用するには、熱赤外センサ開発が今後進み、より安価でより高分解能のセンサが開発されるか、低分解能のセンサを採用し低解像度の熱赤外データに対応するようなジェスチャー認識ソフトウェアを開発する必要がある。現在、低分解能(8×8)ではあるが比較的安価に入手することが可能な熱赤外センサも存在しており、そのようなセンサの応用が実用上期待できる。

また、熱赤外画像ではなく可視画像を対象とした研究であるが、低解像度のナンバープレート画像のパターン認識については既に研究がなされている[11]。本研究における低解像度画像に対する認識シミュレーションはこれを発展させたものである。

1.3 本研究の目的

本研究は、安価な低分解能熱赤外センサを信号の観測ハードウェアとし、ジェスチャー認識を行う非接触ユーザ・インターフェースシステム開発に関するものであり、その第一段階の基礎研究として、認識アルゴリズムの検討を行った。

手のひらのジェスチャー(今回は静止パターン)を対象とし、熱赤外センサを用い取得した温度情報を元に作成した 8×8、16×16 の低解像度画像をデータとし、それらの低解像度画像を用いて認識シミュレーションを行う。低解像度空間においてもパターン識別が可能であることを示すことで以下の二つの要求が満たされることを示したい。

- 1) 比較的安価で小型の入力ハードウェアで非接触ユーザ・インターフェースを構成する。
- 2) 低解像度情報に対応する。

様々なシーンで熱赤外センサを用いた非接触ユーザ・インターフェースを活用するためには、1)の要求は非常に重要である。製品として成立されるためには美観を損なうような仕様は敬遠されるため、小型な機器で構成出来ることは好ましい。また、高精細な熱赤外センサは非常に高価であるが、低分解能のセンサならば比較的安価に用意することができる。

将来的にハードウェアの開発が進み安価に高分解能熱赤外センサを用意することができるようになった場合でも、2) 要求を満たすことで更に汎用的に情報を扱うことが可能となると考えられる。例えば、比較的高解像度な情報を取得することが可能であったとしても、欲しい情報がその情報の一部であった場合、必然的に低解像度情報となる。低解像度画像に対するパターン認識が可能であれば、そういった状況にも対応することが出来る。

パターン認識においてどの程度の画像サイズが必要であるかを確認するため、8×8だけでなく、16×16 に関しても認識シミュレーションを行う。

第 2 章

熱赤外画像のパターン認識手法の原理

本研究は低解像度画像のパターン認識のために、図 2.1 のように大きく分けて 2 つの要素を検討する。

まず、第一に熱赤外画像から必要となる手ひらの抽出方法の検討である。不必要な情報を適切に除外し、観測ノイズを低減することは、安定したパターン認識を行う際に必須である。

そして第二に特徴量に対応する最良なパターン識別アルゴリズムの検討である。本論文では、パターン識別アルゴリズムの選定において「離散コサイン変換(DCT)を利用した認識手法」と「Z 型フーリエ記述子を利用した認識手法」、「P 型フーリエ記述子を利用した認識手法」についてシミュレーションによる評価を行っている。抽出方法と認識アルゴリズムはそれぞれ並行して試験し、それぞれ最も適切な手法、組み合わせを検討する必要がある。図 2.2 は離散コサイン変換を用いたパターン認識に関する評価の流れ、図 2.3 はフーリエ記述子を利用した認識手法に関する評価の流れを示したものである。

それらの手法を検討した上で、熱赤外低解像度ジェスチャー画像に対して安定して判別を行うシステムを開発していく。

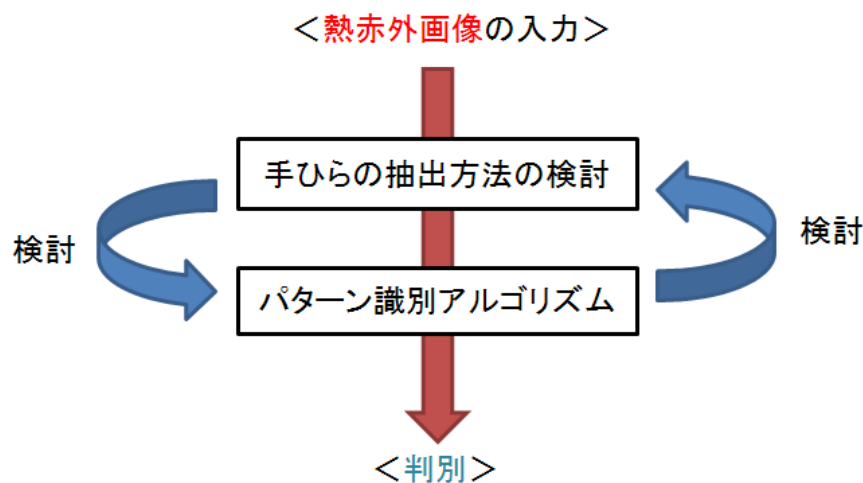


図 2.1 研究要素の説明

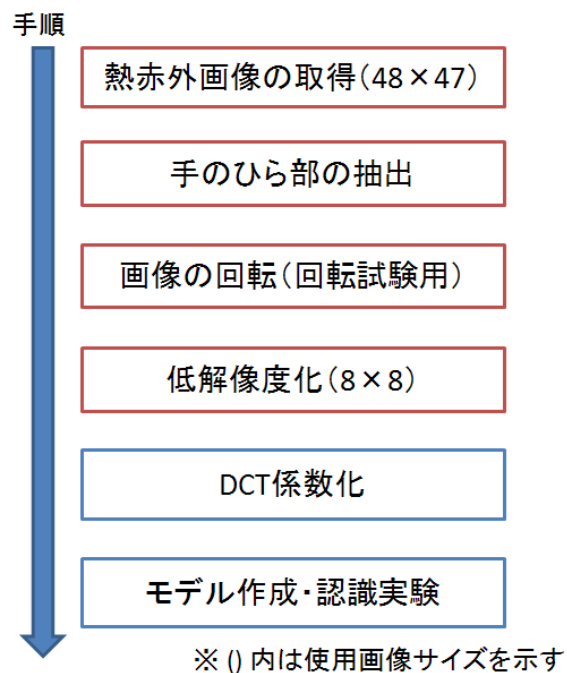


図 2.2 離散コサイン変換(DCT)を利用したパターン認識

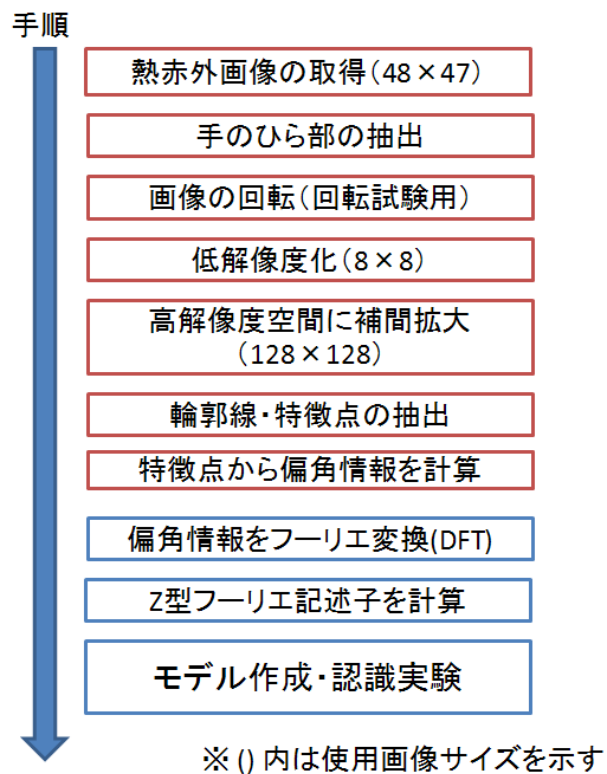


図 2.3 フーリエ記述子を利用したパターン認識

2.1 熱赤外データの前処理

本研究では熱赤外データを使用する。熱赤外データの取得には熱赤外センサ・チノー製「TP-L0225EN」を用い、縦 48×横 47 要素の温度データを取得する。このセンサの温度分解能は 0.1℃である。取得した温度情報は、人間が容易に解釈できる 256 段階のグレースケール画像に変換する。温度データを取得する際、当センサでは測温範囲の上限と下限を設定することが可能である。そして、本研究では人間の温度の取得が目的対象であることから、上限と下限は広くとも±5℃の範囲内で設定を行う。そのため、256 段階の分解能をもつ画像形式であれば十分な温度分解能で温度情報を保存することが可能である。

データの取得の際、センサとジェスチャーを表現する手のひらとの距離は常に 30cm を維持する。取得した温度データは、図 2.4 のような CSV ファイル形式で出力される。またこれを画像に変換したものを図 2.5 に示す。(図 2.5 は図 2.4 のデータをもとに作成された画像である)

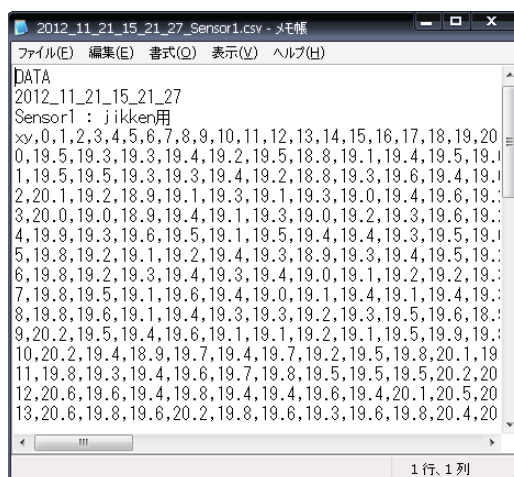


図 2.4 取得温度データ

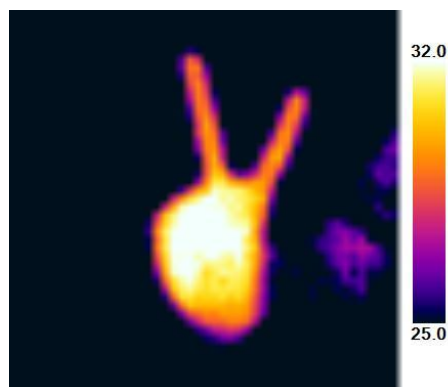


図 2.5 温度データの画像表現

2.2 ジェスチャーの抽出

手のひらのジェスチャーを抽出する際、パターン認識においてノイズになると予想される手のひら以外の不要部分を取り除くため、クリッピングを行うことにする。この際、単純に人体部分の抽出に適した閾値を設定して2値化しただけでは、首や顔など想定外の人体部位が除外できず、画像にノイズとして残る可能性がある。これを防ぐために、本研究では次の手順でクリッピングを行う。この手順の視覚的な表現は図 2.6 で示す。

1. 適切と思われる閾値 k を設定する。
2. 画像内の手のひらの温度と思われる値(ピーク値)を持つ画素の座標 $[i_0, j_0]$ を求める。
3. 手順 2.で得られた画素 $[i_0, j_0]$ の8近傍を走査していく。
4. 走査した画素 $[i_0, j_0]$ の温度が閾値以上ならば、ラベル 1、未満ならばラベル 0 を割り当てる。
5. 手順 4.でラベル 1 を割り振られた画素 $[i_1, j_1]$ の8近傍を走査していく。
6. 手順 5.でラベル 1 を割り振られた各画素に対して手順 4.手順 5.の処理をラベル 1 が見つからなくなるまで繰り返す。
7. 繰り返し処理後ラベル 1 を持つ画素のみを抽出する。

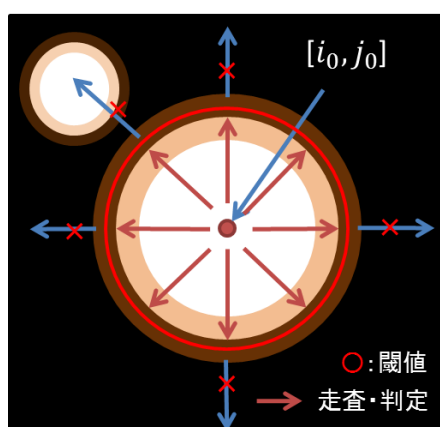


図 2.6 クリッピング手順の視覚的説明

この手順において適切な閾値 k の値の選択は非常に重要である。閾値が適切値より大き過ぎれば当然手のひら以外の部分もまた抽出され、また適切値より小さ過ぎれば必要な形状を抽出する前に処理が終了するからである。今回の研究では、閾値を 0~255 レベル間を 51 段階(レベル幅 5)で変更したサンプル画像をそれぞれ作成し、最も適当な値を主観的に選択した。この過程によって作成したサンプル画像の一部を図 2.7 に示す。

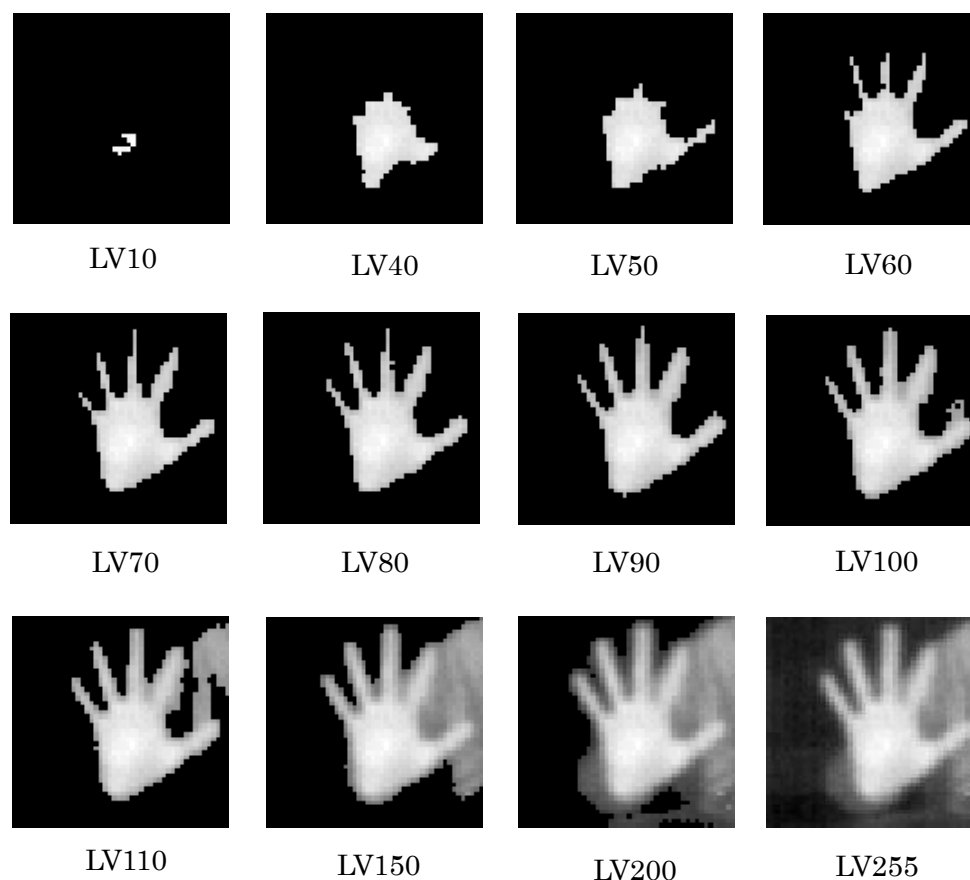


図 2.7 抽出サンプル画像(値は閾値)

図 2.7 のサンプルから、手のひら部分のみを最も適切に抽出できているのは閾値がレベル 90 程度のときであると判断した。なお、この処理において残した部分に関しては 2 値化などの処理は行なっていない。これは今後の判別処理において手のひら内のレベルの勾配を考慮する可能性を失わないようにするためである。パターン認識・判別手法の選択次第では適宜 2 値化などの処理を行うが、汎用的に利用するための画像のオリジナルデータとしてはそれらの処理は行わない。

実際の非接触ユーザ・インターフェースの開発において、この閾値の決定方法は非常に重要な問題になると予想される。今回はシミュレーションの簡便化のために主観的に閾値を決定しているが、非接触ユーザ・インターフェースとしてシステム運用する際は、自動で閾値を決定するか、常に一定にした閾値で対応できるように温度情報を補正する必要がある。本論文では検討しないが、それらの補正方法として空間の周辺温度情報を別のセンサで取得しておき、事前に登録しておいたパーソナル平均体温情報と比較し、閾値を決定する方法などが考えられる。

2.3 離散コサイン変換を用いたパターン認識

2.3.1 離散コサイン変換

コサイン関数だけを用いる離散コサイン変換(以下 DCT)によって得られる DCT 係数を特徴ベクトルとして利用するパターン認識手法[1]。DCT の基底関数としてはいくつかの種類があるが、今回の認識手法では次式の DCT 基底関数を用いる。

$$\varphi_k[i] = \begin{cases} \frac{1}{N} & (k = 0) \\ \sqrt{\frac{2}{N}} \cos \frac{(2i+1)k\pi}{2N} & (k = 1, 2, \dots, N-1) \end{cases} \quad (2.1)$$

これら横方向の基底関数 $\varphi_k[i]$ 、縦方向の基底関数 $\varphi_l[j]$ から 2 次元 DCT 関数 $\varphi_k[i]\varphi_l[j]$ を求め、これを 2 次元 DCT 基底として使用する。

2 次元の DCT 係数を $F_{k,l}$ とし、サイズ $N \times N$ の 2 次元画像信号を $f_{i,j}$ に対する離散コサイン変換は次式(2.2)で表される。

$$F[k, l] = \sum_{j=0}^{N-1} \sum_{i=0}^{N-1} f_{i,j} \varphi_k[i] \varphi_l[j] \quad (2.2)$$

ある画像濃度値 $f[m, n]$ に対する DCT 係数を $F[k, l]$ とすると振幅スペクトルは $F[0,0]$ に近い係数であるほど値が大きい (図 2.8)。また、 $F[0,0]$ は直流成分(DC 係数)であり、その他の $F[k, l]$ は交流成分(AC 係数)である。通常の写真は直流成分が最も大きく、高周波成分ほど小さくなる。

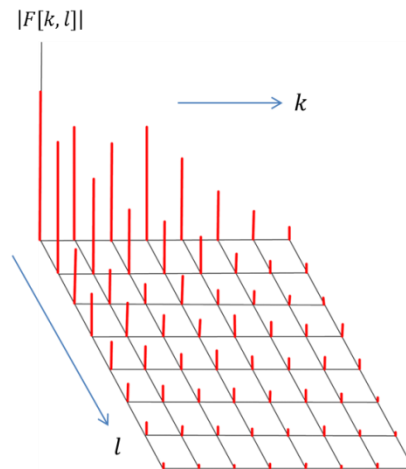


図 2.8 DCT 係数のスペクトルの大きさ

2.3.2 認識システムの流れ

DCT によるパターン認識では DCT 係数 $F_{k,l}$ がそれぞれのジェスチャーによって値が変化することを利用し、モデルの作成および判別を行う。DCT を用いたパターン認識システムにおける処理の流れを図 2.9 に示す。

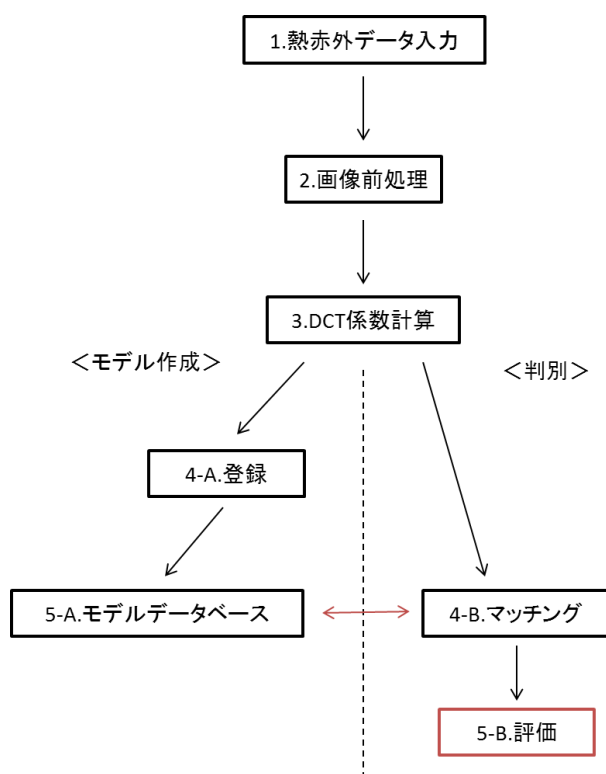


図 2.9 認識システムにおける処理の流れ

手順 1 と手順 2 において、レベル 256 段階の熱赤外データを使用し、手のひら部分を抽出する。手順 3 で画像に対して DCT 係数の計算を行うが、画像サイズが常に $N \times N$ で一定であるとする式(2.1)の 2 次元 DCT 基底はテーブルで用意しておけば良く、繰り返し計算を削減することができる。2 次元 DCT 基底に対し画像信号を入力し DCT 係数を計算する。

判別の際、ラベル付けをした事前にモデルデータを用意し、データベース化しておく。ここでモデルデータベース作成に関していくつかの注意点がある。まず、用意するモデルデータは必要最低限の数であることが好ましい。モデル数が多くなればなるほど、マッチングの際に評価回数が多くなるためである。しかし、モデル数の不足によって誤判別が多発するようならば適宜モデルを追加していく必要がある。特にジェスチャーの傾きによって特異な形状になる場合は、基本的に特異形状用モデルを用意しておくことで対策しておくのが無難である。

2.3.3 評価式

画像サイズを $N \times N$ とすると得られる DCT 係数も $N \times N$ 個である。基本的には低周波側($F[0,0]$ 側)の成分を用いてマッチングを行う。これは、低周波成分であるほど振幅が大きく、各認識対象間の識別への寄与が大きいためである。マッチング処理においては、モデル DCT 係数 $C_m[i] = F_{k,l}$ と入力画像の DCT 係数 $C_{\text{input}}[i] = F_{k,l}$ とのユークリッド距離 d_e を評価する。評価式は以下の式となる。

$$d_e = \sqrt{\sum_{i=1}^{n-1} (C_m[i] - C_{\text{input}}[i])^2} \quad (2.3)$$

低周波側の $M \times M$ 個の DCT 係数を評価に用いる。 M の大きさは、大きければ良いとは限らず、認識対象とモデルデータの関係から適宜選択する必要がある。このとき、 $n = M \times M$ であり、 $i = k + lM$ である。この評価式のユークリッド距離 d_e が小さいほどモデルとの近いデータということになる。この認識手法では、各点の位置関係のみを評価するので分散等は考慮しない。

モデルデータと認識対象画像との DCT 係数値の違いを図 2.10 に例として示す。

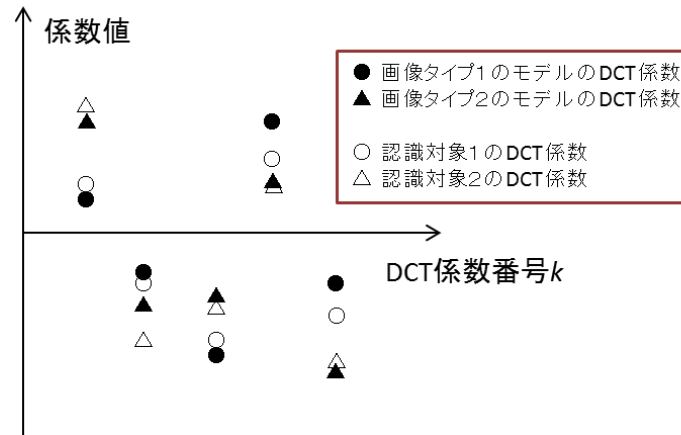


図 2.10 DCT 係数の違いのグラフによる表現

2.4 フーリエ記述子を用いたパターン認識

2.4.1 フーリエ記述子の概要

2次元の平面状の図形であっても輪郭の点座標や偏角をある固定点からの相対的な位置の関数とみなすと1次元の波形となる。また、この波形をフーリエ級数展開して得られた有限個のフーリエ係数を用いて線形図の特徴を表現することが可能であり、このような表現方法をフーリエ記述子と呼ぶ[1]。

通常の波形は高周波成分を取り除き、いくつかの低周波成分のフーリエ記述子を用いることで波形を良い近似で再生することが可能であり、フーリエ記述子は線形図の特徴をよく表現するであろうことが期待される。また、高周波成分の情報を除き、低周波数成分の係数のみで形状再生が十分可能であることから、評価する際、高速で評価することが可能である。これは、機器などに組み込む際に非常に重要である。

フーリエ記述子にはいくつかの種類があり、その代表的なものをいかに簡単に説明する。

(i) 位置座標表現：G型フーリエ記述子

対象を閉曲線とし、曲線上の位置座標を以下の式を用いて複素平面で表す。

$$z(s) = x(s) + jy(s) \quad (2.4)$$

ここで $x(s)$ 、 $y(s)$ は閉曲線上の点の x 座標と y 座標を表す関数とする。変数 s は始点からの曲線の長さとする。閉曲線の場合は常に始点と終点が一致し、多角形に対しても不連続点が生じない周期的な連続関数となる。(2.4)式は連続した信号について表したものであり、 $z(s)$ を等間隔にサンプリングして離散データ z_i として表現した場合、離散フーリエ変換は以下の式で表される。

$$C_k^{(G)} = \frac{1}{N} \sum_{i=0}^{N-1} z_i e^{-j\frac{2\pi}{N}ki} \quad (2.5)$$

(2.5)式で与えられる $C_k^{(G)}$ がG型フーリエ記述子であり、これを用いて図形を表現できる。開曲線に対して行くと、始点と終点で不連続な関数となるため、高次成分が無視できない大きさとなる。これにより低い周波数係数だけでマッチングの判定を行うことが難しくなる。

(ii) 偏角関数表現：Z型フーリエ記述子

曲線上のある位置(0)を始点とし、始点からの距離 s の位置における曲線接線と水平軸とのなす角を $\theta(s)$ とする。始点から反時計回りに線上を移動したとすると、 $\theta(s)$ は図 2.11 のように定義される。曲線の全長を L とすると、 $\theta(0)$ と 1 周して始点へと至った場合の $\theta(L)$ との間では、 $\theta(L) = \theta(0) + 2\pi$ となり不連続点が生じる。不連続点が存在すると高次成分が大きくなり、好ましくない。不連続点の発生を防ぐため、以下のような正規化偏角関数を用いる。

$$\theta_n(s) = \theta(s) - \theta(0) - 2\pi \frac{s}{L} \quad (2.6)$$

また、一定の間隔で曲線をサンプリングした離散データを使用した場合の正規化偏角関数は以下のように定義される。

$$\varphi_i = \theta_i - \theta_0 - \frac{2\pi}{N} i \quad (i = 0, 1, 2, \dots, N-1) \quad (2.7)$$

この正規化偏角関数を離散フーリエ変換すると、

$$C_k^{(Z)} = \frac{1}{N} \sum_{i=0}^{N-1} \varphi_i e^{-j \frac{2\pi}{N} k i} \quad (2.8)$$

となる。これにより得られた $C_k^{(Z)}$ を Z 型フーリエ記述子と呼ぶ。この記述子は閉曲線のみを使用できるという制約があるものの、平行移動、拡大・縮小に不変であるという性質をもつ。

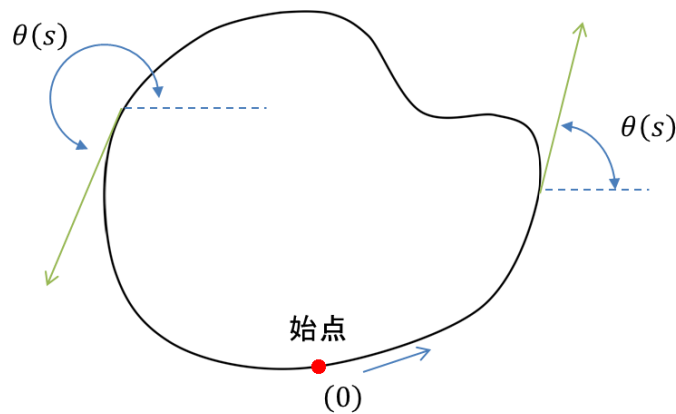


図 2.11 接線と水平線のなす角(偏角)の表現

(iii) 指数関数表現：P 型フーリエ記述子

曲線を図 2.12 のように折れ線近似した場合、偏角 θ_i の指数関数と複素平面における座標 z_i との関係は次式ようになる。

$$\begin{aligned} w_i &= \exp(j\theta_i) = \cos\theta_i + j\sin\theta_i \\ &= \frac{x_{i+1} - x_i}{\delta} + j \frac{y_{i+1} - y_i}{\delta} = \frac{z_{i+1} - z_i}{\delta} \end{aligned} \quad (2.9)$$

この式における δ は線分の長さであり、

$$\delta = |z_{i+1} - z_i| \quad (2.10)$$

で表される。(2.9)式で表される w_i を周期関数として離散フーリエ変換を求め、

$$C_k^{(P)} = \frac{1}{N} \sum_{i=0}^{N-1} w_i e^{-j\frac{2\pi}{N}ki} \quad (2.11)$$

を得る。これにより得られた $C_k^{(P)}$ を P 型フーリエ記述子と呼ぶ。式(2.9)で表現された周期関数 w_i の定義から予想されるように、P 型フーリエ記述子は平行移動、拡大・縮小に対して不変であり、画像の回転に対しても絶対値は不変である。Z 型フーリエ記述子とは異なり、閉曲線だけでなく開曲線に対しても有効であるという特徴がある。

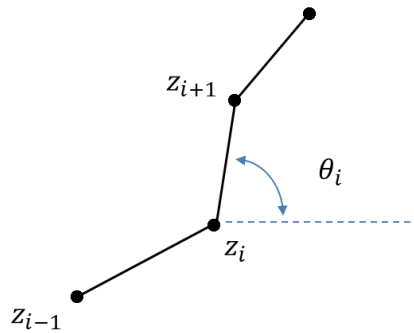


図 2.12 曲線の折れ線近似

2.4.2 輪郭抽出

図形の形状をフーリエ記述子によって表現する場合において、閉曲線又は開曲線の輪郭のベクトルデータを抽出する必要がある(図 2.13) (輪郭抽出手法は文献[1]を参照) また、文献[10]より)。また、抽出される輪郭は 8 近傍接続によるものとする。二値化データでは無く、輝度勾配を含んだ画像に対して輪郭追跡を行う場合、輝度 $LEVEL \neq 0$ を基準とする方法と、閾値 LK を事前に設定し輝度 $LEVEL \geq LK$ を基準とする方法があり、いずれかの判定基準に基づいて輪郭追跡を行う。本研究では図 2.14 のような画像を用いるので、閾値 LK を設定する方法が好ましい。その理由は、形状抽出において可能な限り輝度の勾配情報が利用できるからである。ただし、閾値が大き過ぎた場合抽出した形状が小さくなり、閾値が小さ過ぎた場合には形状が大き過ぎる抽出され過ぎることがあるため、閾値を適切に選択する必要がある。

ここからは抽出の具体的な抽出の流れを説明する。

追跡処理を開始する前に、追跡対象となる画像 f_{ij} に対して、同じ行数と列数の配列 S_{ij} 、 F_{ij} を用意する。配列 S_{ij} は、各座標がラスタ走査済みであるか否かを判定するために使用し、配列中の値は全て 0 で初期化しておく。また、 F_{ij} は出力される輪郭データであり、この配列も 0 で初期化しておく。

- (1) 座標(0,0)から閾値 LK 以上の値をもつ画素が現れるまでラスタ走査していき、その座標における S_{ij} の値が 0 であった場合、それを始点とする。その際、始点とした座標を含めてラスタ走査が済んだ座標における S_{ij} に 1 を代入する。
- (2) 手順(1)で始点とした座標の 8 近傍を反時計回りで輝度 $LEVEL \geq LK$ であるかの判定を行う。この手順で輝度 $LEVEL \geq LK$ の条件を満たす座標が無かった場合、調査点は孤立点であるので手順(1)に戻る。また、この判定処理を行った座標における S_{ij} にも 1 を代入し、調査済みであることを記録しておく。孤立点で無かった場合、その座標における F_{ij} に 255 を代入する。
- (3) 手順(2)における判定基準を最初に満たした座標を新たな追跡点とする。その座標における F_{ij} に 255 を代入すると共に、その追跡点に対して再び反時計回りで輝度 $LEVEL \geq LK$ の条件を満たすかどうか、またその座標における S_{ij} が 0 であるかの判定を行う。手順(2)と同様に、調査済みの座標に対しては S_{ij} に 1 を代入する。
- (4) 手順(3)の処理を繰り返し、反時計回りの判定処理中で S_{ij} の値が 1 しかなかった場合、また追跡点と追跡始点が同じ座標に成ったら追跡終了とする。ほかの追跡点を必要とする場合は、手順(1)へと戻る。

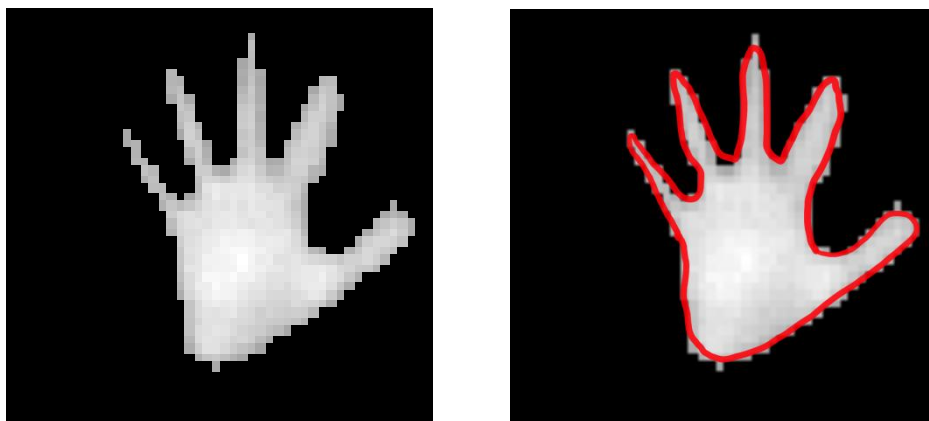


図 2.13 輪郭の抽出

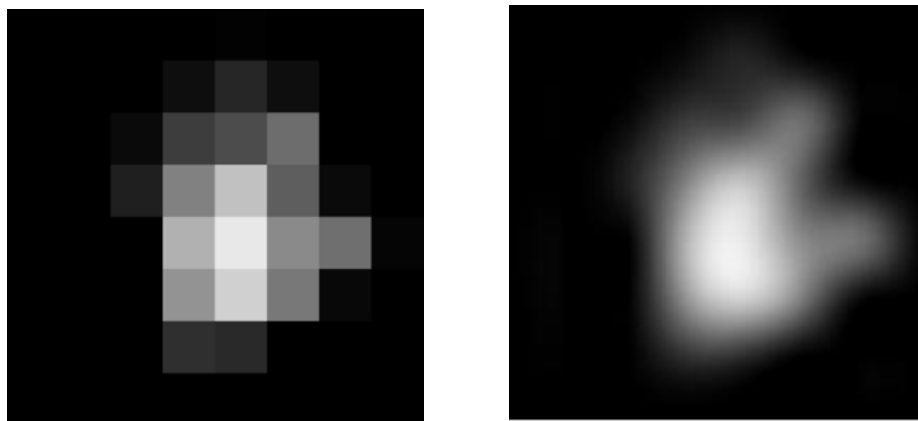


図 2.14 低解像度熱赤外画像

左：8×8 ピクセル画像、右：128×128 に補間処理した画像

2.4.3 代表点の抽出

輪郭抽出を行い輪郭の特徴ベクトルを用意した後、通常、特徴量を保存しつつデータ数を少なくする処理を行う。曲線をいくつかの線分で近似し、線分と線分との接点を代表点とする。代表点を求める処理の基礎となる3つの手法を説明する。

(i) 距離(画素)単位を利用する方法

一つ目の手法として、1本の曲線に対して一定間隔で代表点を配置し分割していく手法である(図2.15)。間隔を距離 l または調査画素数 p で設定し、始点から輪郭を追跡していく。この手法において得られる代表点の数 N は、輪郭の全長または全画素数によって変わってくるため、一定数の代表点を用いた認識アルゴリズムに適用する際は、他の手法と組み合わせ、代表点数を間引くことによって適宜対応する必要がある。また、始点と終点との距離が、距離 l を常に保つとは限らない。

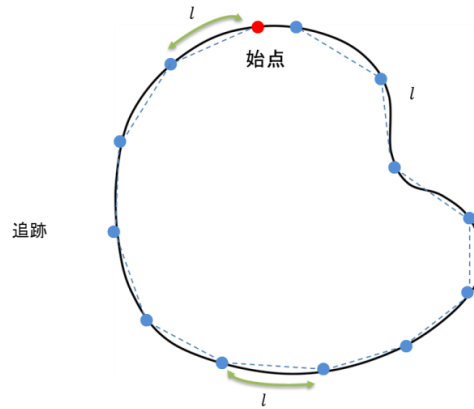


図 2.15 輪郭に対する距離を利用する方法

(ii) 輪郭の全長を代表点数 N で等間隔に分割する方法

二つ目の手法として、曲線の全長を計算しそれを N 個の代表点で分割するものである。図2.16のように代表点間の曲線上の距離を l 、曲線の全長を L とすると、

$$l = \frac{L}{N} \quad l \geq |P_i P_{i+1}| \quad (i = 0, 1, 2, \dots, N-1) \quad (2.12)$$

である。また、閉曲線では始点と終点が一致する。(i)の手法と異なり代表点数を一定とすることが出来るが、図形によって全長が異なるため、異なる図形間では距離 l の大きさが異なる。また、今回のシミュレーションにおける全長の長さは、1画素を単位とし、輪郭曲線上を追跡した際の累積画素数としている。斜め方向に関しては横や縦方向と同様に距離1画素分と近似しているため、実際の全長と差が生じてしまうが、許容範囲内の誤差であると考ええる。また、全長計算において輪郭を追跡

する際、追跡画素における 8 近傍に複数の画素があった場合、4 近傍中にある画素を優先的に追跡していくこととする。

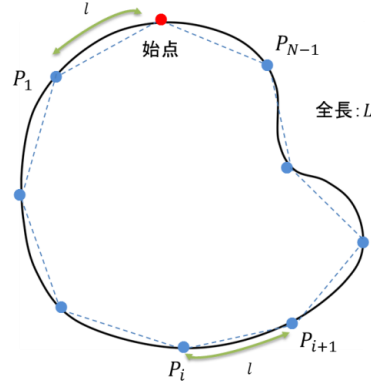


図 2.16 輪郭の全長を代表点数 N で分割する方法

(iii) 曲率を利用する方法

三つ目の手法として代表点抽出に曲率を利用したものが挙げられる。曲率が大きく変化する点を折れ点とする。図 2.17 が示すように折れ点の計算には 3 点を事前にサンプリングする必要がある、手法(i)もしくは手法(ii)を利用して仮の代表点 P_{i-1} 、 P_i 、 P_{i+1} を抽出しておく。これから、ある間隔をもった仮の代表点どうしでそのなす角度の差を求め、設定した閾値 ϕ 以上ならば折れ点とし代表点とする。2 直線の角度 θ は以下の式で求める。

$$\begin{aligned} \cos\theta &= \frac{\langle a, b \rangle}{\|a\| \cdot \|b\|} \\ &= \frac{(x_1 - x_0)(x_2 - x_1) + (y_1 - y_0)(y_2 - y_1)}{\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} \end{aligned} \quad (2.13)$$

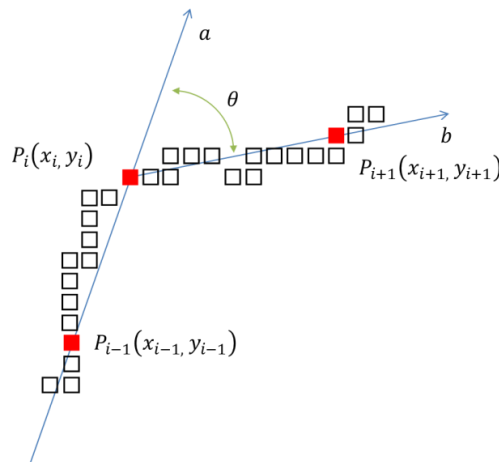


図 2.17 2 次元ベクトルのなす角

2.4.4 評価式

フーリエ記述子のモデル係数 $C_k^{(\text{model})}$ と入力係数 $C_k^{(\text{input})}$ のマッチングを行う。画像の輪郭を表現する記述子を周期関数を f_i とし、これを離散フーリエ変換する。

$$C_k = \frac{1}{N} \sum_{i=0}^{N-1} f_i e^{-j\frac{2\pi}{N}ki} \quad (2.14)$$

また、この式に対してオイラーの公式を適用すると以下の式になる。

$$C_k = \frac{1}{N} \sum_{i=0}^{N-1} f_i \cos\left(\frac{2\pi ki}{N}\right) - j \frac{1}{N} \sum_{i=0}^{N-1} f_i \sin\left(\frac{2\pi ki}{N}\right) \quad (2.15)$$

記述子として本研究で利用する Z 型、P 型を用いると次のフーリエ記述子が得られる。まず Z 型フーリエ記述子は、

$$C_k^{(z)} = \frac{1}{N} \sum_{i=0}^{N-1} \left\{ \left(\theta_i - \theta_0 - \frac{2\pi}{N}i \right) \cos\left(\frac{2\pi ki}{N}\right) \right\} - j \frac{1}{N} \sum_{i=0}^{N-1} \left\{ \left(\theta_i - \theta_0 - \frac{2\pi}{N}i \right) \sin\left(\frac{2\pi ki}{N}\right) \right\} \quad (2.16)$$

となり、P 型記述子は、

$$C_k^{(p)} = \frac{1}{N} \sum_{i=0}^{N-1} \left\{ \left(\frac{x_{i+1} - x_i}{\delta} \right) \cos\left(\frac{2\pi ki}{N}\right) - \left(\frac{y_{i+1} - y_i}{\delta} \right) \sin\left(\frac{2\pi ki}{N}\right) \right\} \\ + j \frac{1}{N} \sum_{i=0}^{N-1} \left\{ \left(\frac{x_{i+1} - x_i}{\delta} \right) \sin\left(\frac{2\pi ki}{N}\right) + \left(\frac{y_{i+1} - y_i}{\delta} \right) \cos\left(\frac{2\pi ki}{N}\right) \right\} \quad (2.17)$$

となる。マッチングの特徴ベクトルとしては高次成分を取り除き、また、 C_0 は直流成分を表し、形状の情報を持たないので、 $k = 1 \sim K$ と $k = (N - K) \sim (N - 1)$ の C_k を用いる。式(2.16)、(2.17)を実部と虚部があるので合計 $4K$ 個の成分をもつ特徴ベクトルが得られる。 $C_k^{(\text{model})}$ の特徴ベクトルを a_i 、 $C_k^{(\text{input})}$ の特徴ベクトルを b_i とすると、評価の尺度としてモデルベクトルと入力ベクトルとのユークリッド距離 d_e を次の式で計算する。

$$d_e = \sqrt{\sum_{i=1}^{n-1} (a_i - b_i)^2} \quad (2.17)$$

このユークリッド距離が小さい程、モデル画像と入力画像との特徴の類似性が高い。

本研究では3種類のモデル画像に対して、入力画像がどの特徴ベクトルに最も近いかを計算し、答えとして最短距離のラベルを返すようにする。

第 3 章

離散コサイン変換を用いたパターン認識シミュレーション

3.1 使用する低解像度熱赤外面像

このシミュレーションで使用する画像は、「グー」「チョキ」「パー」の三種類のジェスチャーの温度情報を熱赤外センサ取得したものである。

シミュレーションでは、「画像の回転」「手の傾き」の二つの変化による DCT 係数の変化について考察を行う。

まず画像回転による DCT 係数の変化を見るために、基本となる画像として図 3.1 の 3 種類の画像を用意する。回転画像を用意する前に、これらの画像に対しジェスチャー抽出処理を用いて不必要部分の輝度値をクリッピングし、図 3.2 の 3 枚の画像に変換する。温度情報のクリッピングでは閾値レベルを 90 と設定する。これらの画像を回転角度 0° とし、反時計回りで 10° 毎に 90° まで回転した画像 9 枚を回転によって作成する。したがって、基本画像 3 枚 + 回転画像 27 枚の計 30 枚の画像が生成される(図 3.3)。これらの画像に対するクリッピング処理は、熱赤外センサによるジェスチャー取得の際に行われる処理の一部として模擬するため、 48×47 画素の高解像度画像空間で処理を行う。また、回転処理によって手の長軸部分の情報が失われるのを防ぐために、横 3 画素、縦 4 画素の余白(輝度値 0 空間)を付加した上で回転処理を行う。次に、これら 51×51 画素の 30 枚の画像に対して処理を行い、 8×8 サイズの低解像度熱赤外面像を用意する。縮小処理には線形補間処理アルゴリズムを用いて行う。図 3.4 は 8×8 画素の回転処理を行った低解像度熱赤外面像群であり、回転に関する考察の対象とする。

2 つ目のシミュレーション要素である「手の傾き」に対する画像として。「グー」「チョキ」「パー」の三種類のジェスチャーにおける傾き画像を 12 枚ずつ取得した(図 3.5)。これらの 12 枚の画像に対して縮小処理を行い、 8×8 の低解像度画像計 36 枚を作成する(図 3.6)。傾きに関する考察を行う際は、この図 3.6 の画像群を使用したシミュレーションを行う。



図 3.1 回転シミュレーションにおける「グー」「チョキ」「パー」の基本画像

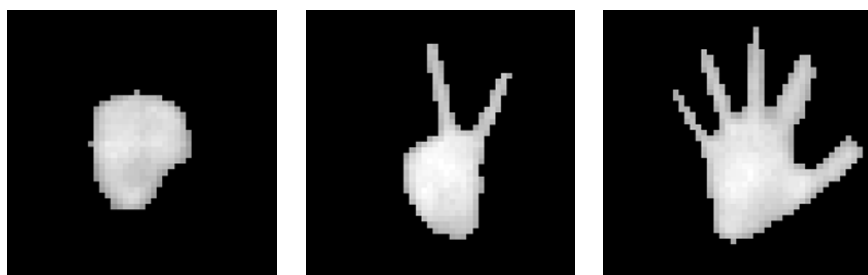


図 3.2 図 3.1 に対してクリッピング処理を行った画像

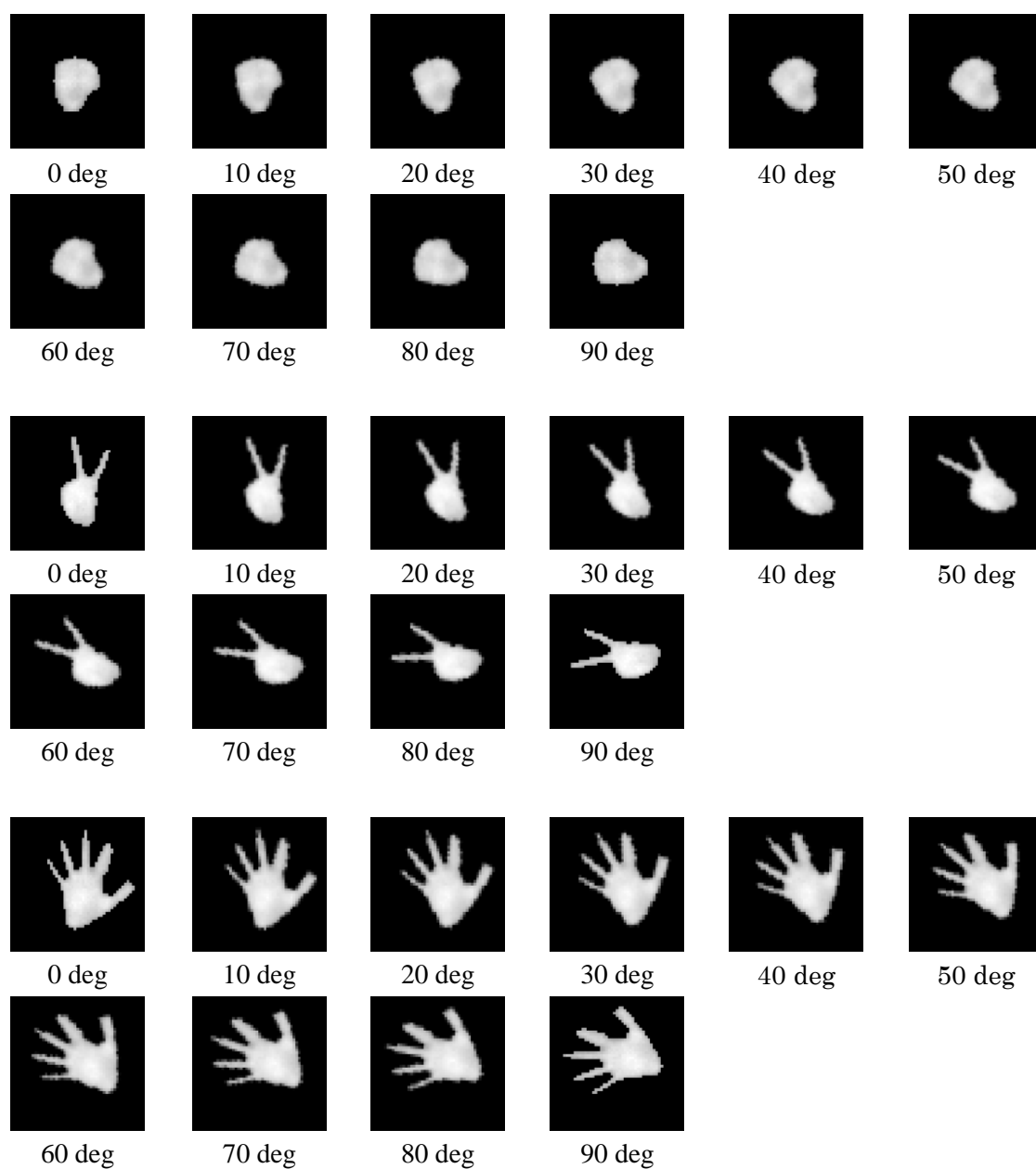


図 3.3 「グー」「チョキ」「パー」の回転画像 30 枚(51×51 画素)

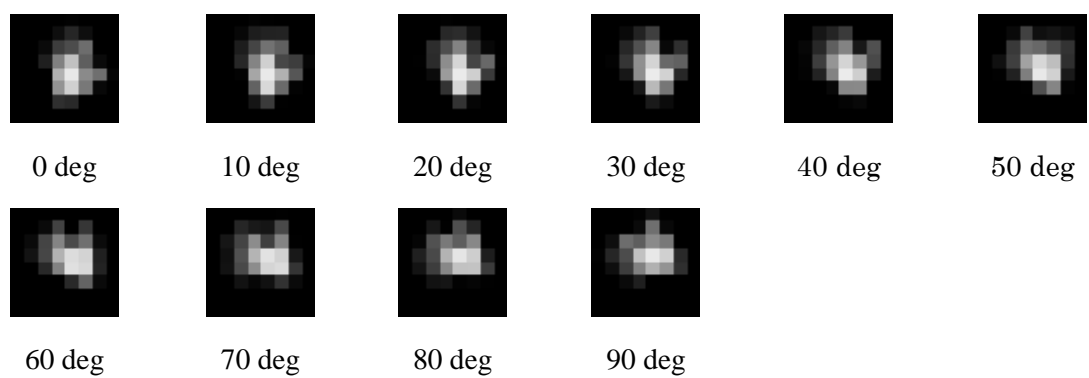
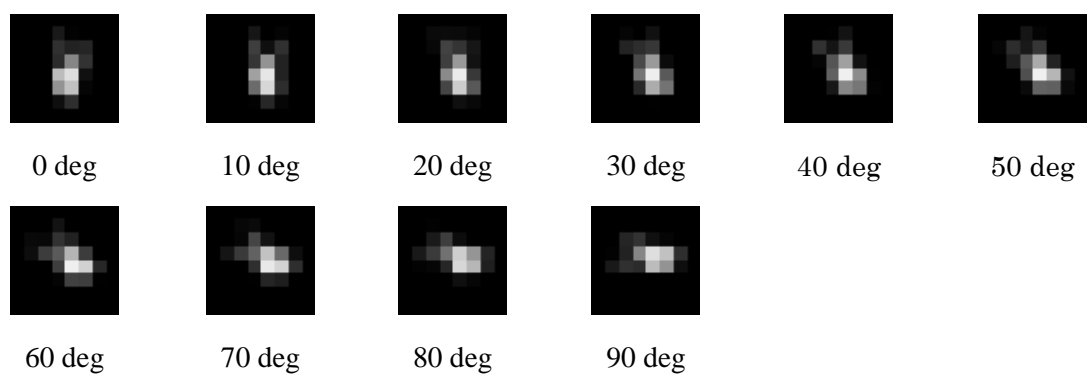
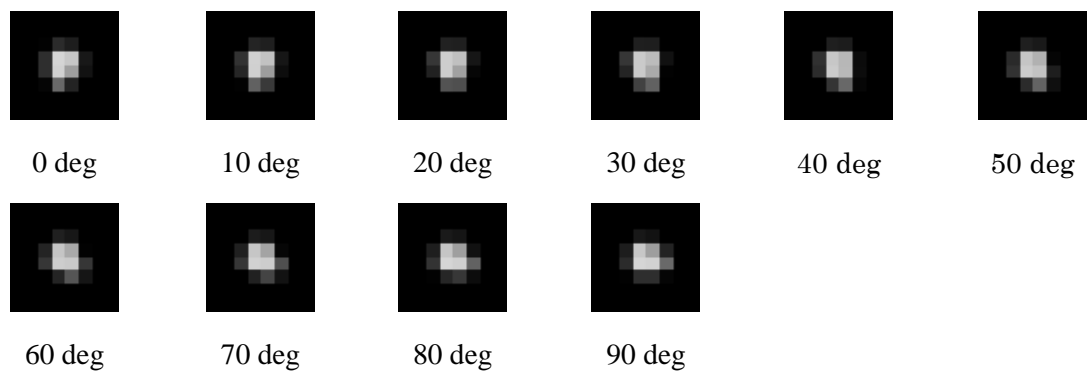
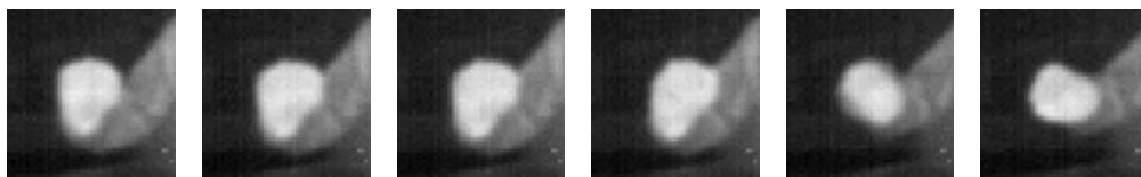


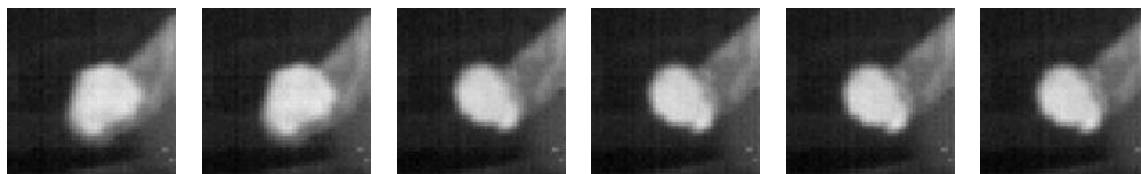
図 3.4 回転シミュレーションに用いる 8×8 画素の低解像度画像

<グー>

横方向の傾き変化

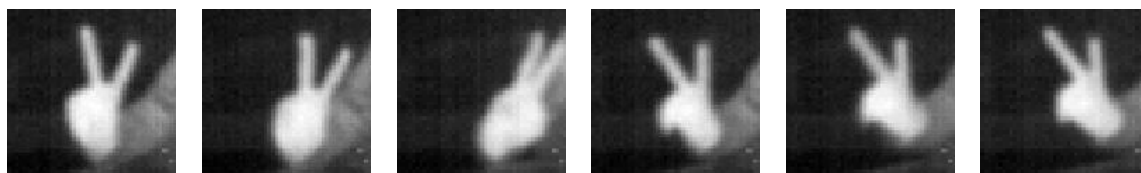


縦方向の傾き変化

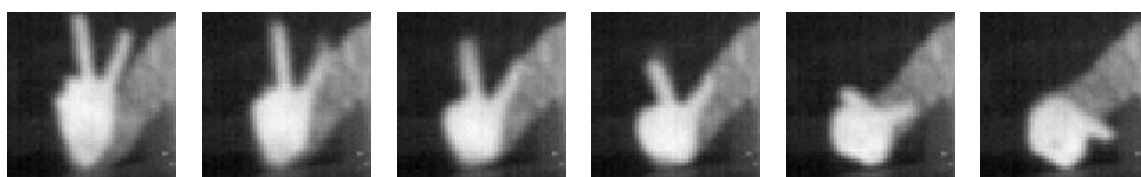


<チョキ>

横方向の傾き変化



縦方向の傾き変化



<パー>

横方向の傾き変化



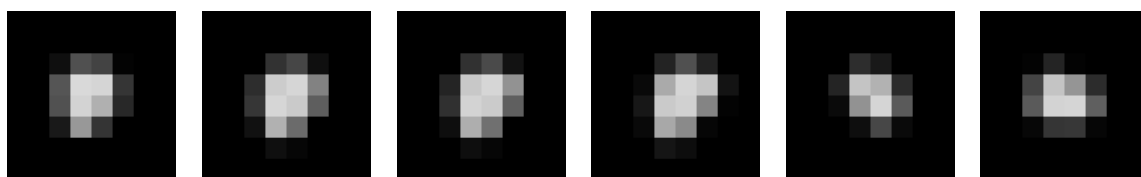
縦方向の傾き変化



図 3.5 傾きによる変化を考察するための熱赤外画像(48×47 画素)

<ゲー>

横方向の傾き変化



1

2

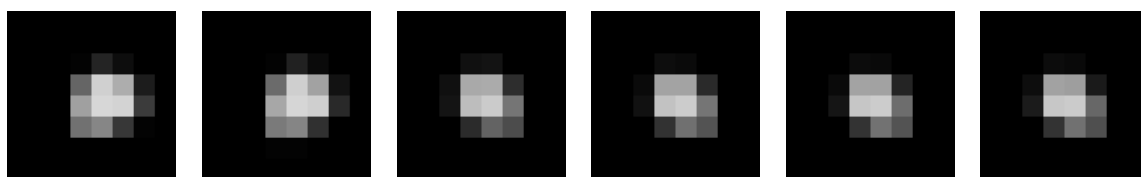
3

4

5

6

縦方向の傾き変化



7

8

9

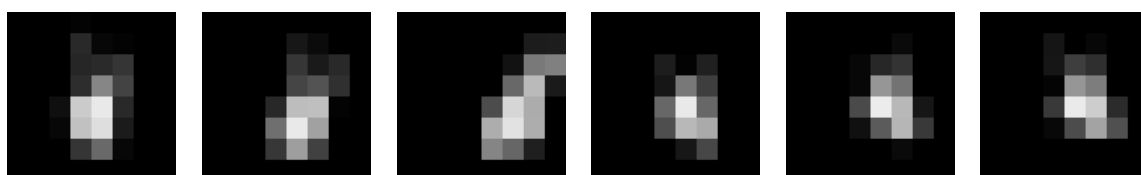
10

11

12

<チョキ>

横方向の傾き変化



1

2

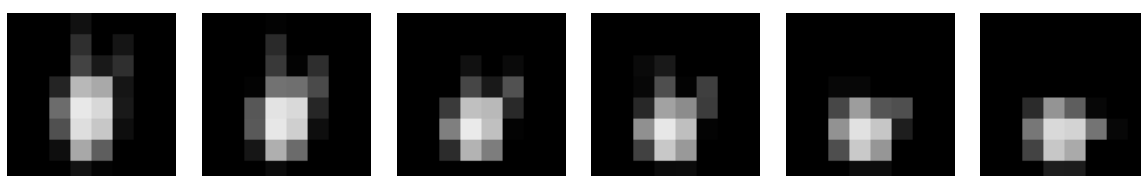
3

4

5

6

縦方向の傾き変化



7

8

9

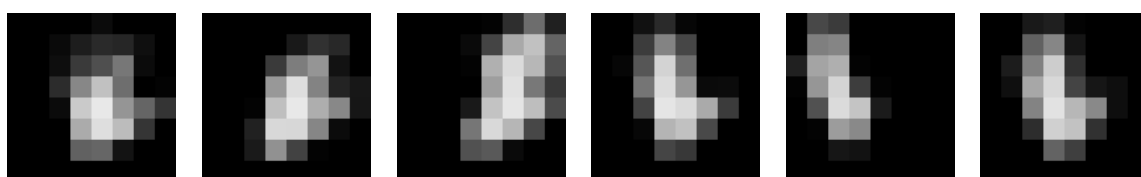
10

11

12

<パー>

横方向の傾き変化



1

2

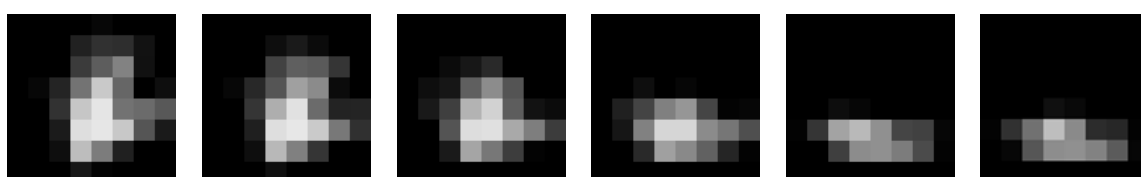
3

4

5

6

縦方向の傾き変化



7

8

9

10

11

12

図 3.6 傾きによる変化を考察するための低解像度熱赤外画像(8×8 画素)

3.2 回転に対する DCT 係数の変化

回転による DCT 係数の変化を見るために図 3.4 の画像群を用いてシミュレーションを行う。図 3.4 の低解像度画像の入力信号のサイズは 8×8 なので、 $N=8$ となる。よって式(2.1)の式は以下のように計算される。

$$\varphi_k[i] = \begin{cases} \frac{1}{8} & (k = 0) \\ \sqrt{\frac{1}{4}} \cos \frac{(2i+1)k\pi}{16} & (k = 1,2,3,4,5,6,7) \end{cases} \quad (3.1)$$

$$\varphi_l[j] = \begin{cases} \frac{1}{8} & (l = 0) \\ \sqrt{\frac{1}{4}} \cos \frac{(2j+1)l\pi}{16} & (l = 1,2,3,4,5,6,7) \end{cases} \quad (3.2)$$

入力信号の座標 (i, j) それぞれの信号に対して式(3.1)、式(3.2)の基底を用いる。

ここで、 $F_{k,l}$ を 2 次元の DCT 係数と呼ぶ。 8×8 画素の 2 次元画像信号 $f_{i,j}$ に対する離散コサイン変換は次式(3.3)で表される。

$$F_{k,l} = \sum_{j=0}^7 \sum_{i=0}^7 f_{i,j} \varphi_k[i] \varphi_l[j] \quad (3.3)$$

8×8 画素の場合、計 64 個の DCT 係数が出力される。次に、実際に回転画像を入力して得られた DCT 係数のグラフを示す。

まず、「ゲー」のジェスチャーに対する DCT 係数の振幅をグラフにしたものを示す。横軸を DCT 係数番号とし、縦軸は係数の振幅を示している。

係数の振幅

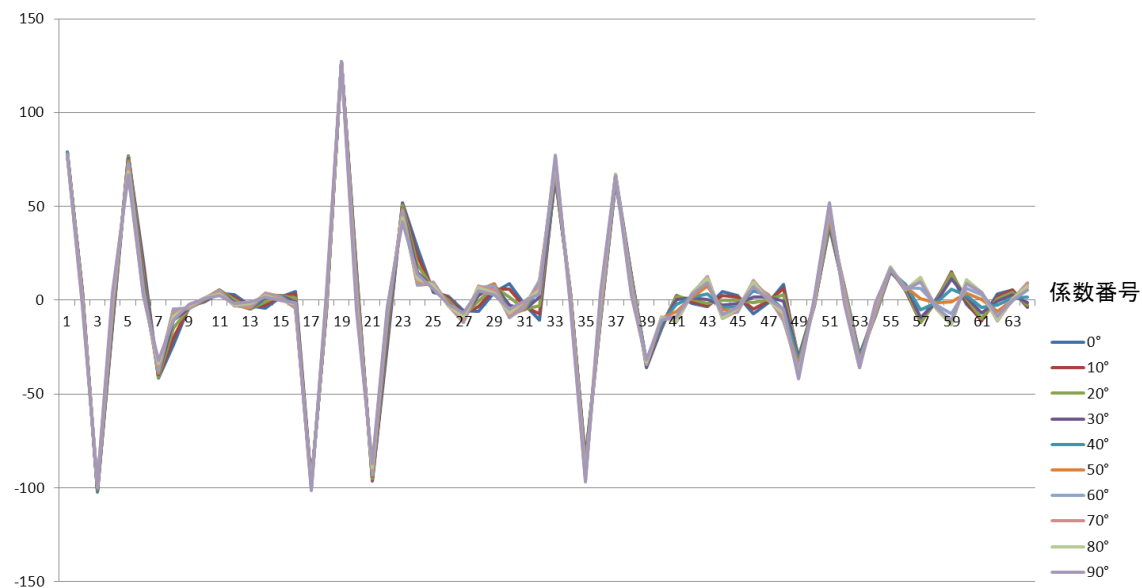


図 3.7.1 <グー>の回転角に対する DCT 係数の変化

係数の振幅

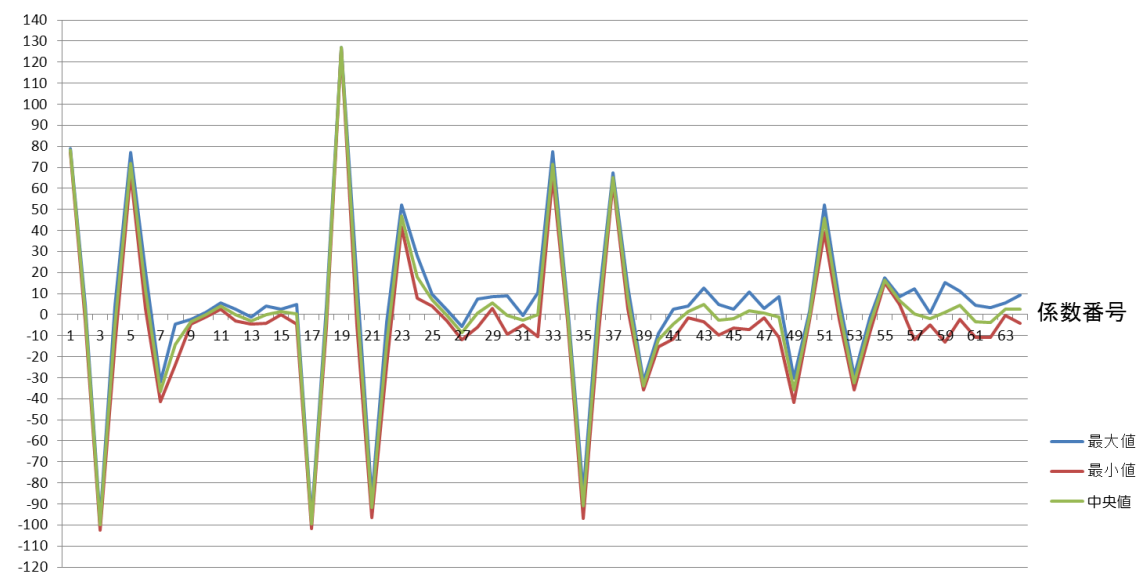


図 3.7.2 <グー>の DCT 係数の変化幅(最大値、最小値、中央値)

「グー」における DCT 係数の振れ幅は基本的に小さく、特に特定の係数で非常に振れ幅が低い。係数 1 番は直流成分なので、画像の輝度値の平均値を示していると考えられる。よって、回転の影響は無いのでほぼ一定の値を示している。

次に、「チョキ」のジェスチャーに対する DCT 係数の振幅をグラフにしたものである。横軸を DCT 係数番号とし、縦軸は係数の振幅を示している。

係数の振幅

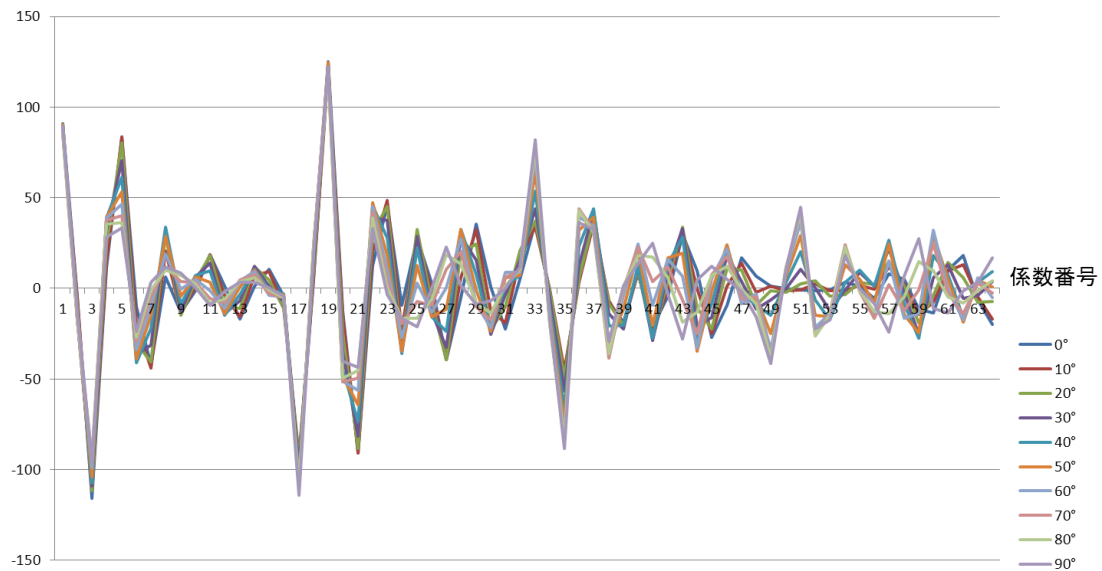


図 3.8.1 <チョキ>の回転角に対する DCT 係数の変化

係数の振幅



図 3.8.2 <チョキ>の DCT 係数の変化幅(最大値、最小値、中央値)

「チョキ」に関する DCT 係数の振れ幅は「グー」に比べて大きい。これは、回転による画像変化が「グー」に比べて大きいことに起因する。しかしながら、特定の係数番号、例えば 10 や 19 は変化が小さい。この番号は回転による変化が小さい周波数であると考えられる。

最後に、「パー」のジェスチャーに対する DCT 係数の振幅をグラフにしたものである。横軸を DCT 係数番号とし、縦軸は係数の振幅を示している。

係数の振幅

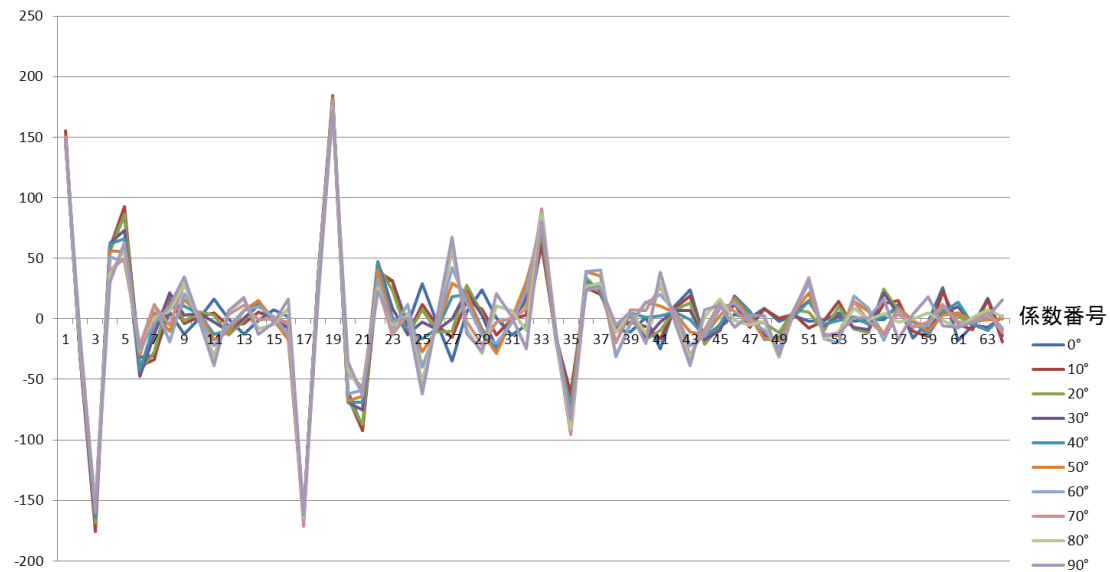


図 3.9.1 <パー>の回転角に対する DCT 係数の変化

係数の振幅

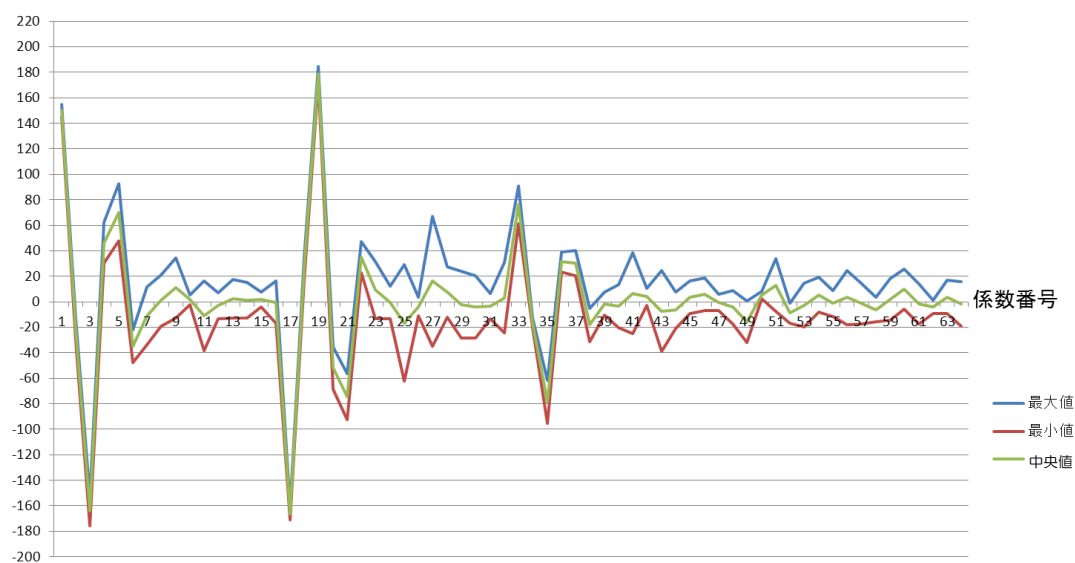


図 3.9.2 <パー>の DCT 係数の変化幅(最大値、最小値、中央値)

「パー」もチョキと同様に「グー」に比べて、回転による DCT 係数の振幅の変化幅は大きい。とはいえ、「チョキ」と同様に特定の係数番号の振れ幅は小さくなっている。「チョキ」と同様の係数番号で振れ幅が小さくなっていることもあり、「チョキ」や「パー」「グー」の判別に利用できる可能性がある。

その可能性を考察するために、振れ幅が比較的小さい係数番号のみを取り出して、ローソクチャートで、各ジェスチャー間における距離を図 3.10 に示した。

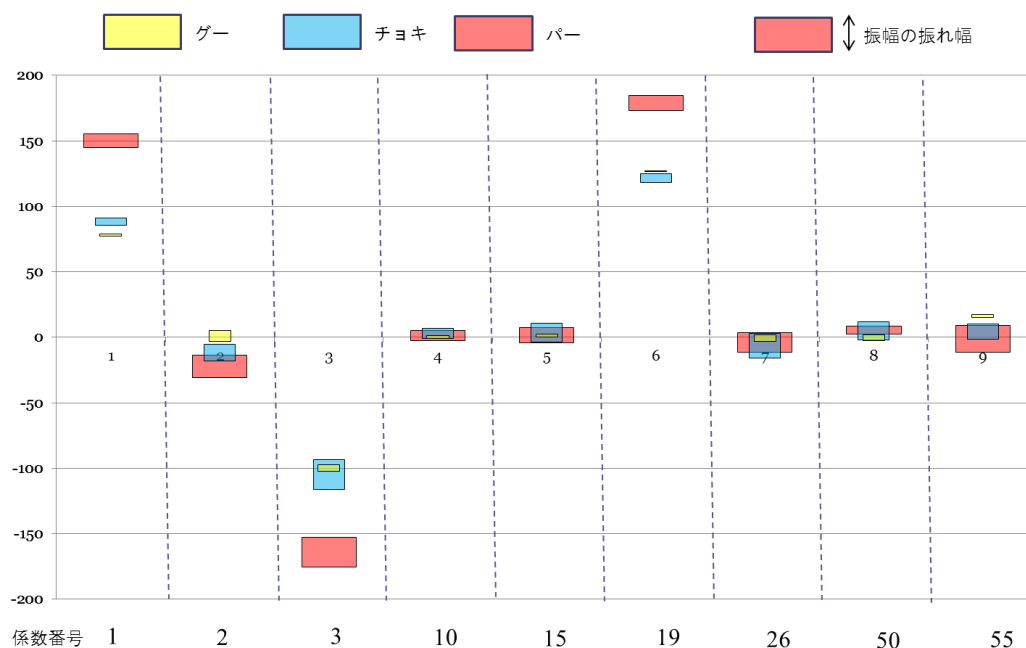


図 3.10 特定の番号における振幅の振れ幅

振幅の振れ幅が互いに重なっている部分もあるが、ジェスチャーによっては分離できているものもある。特に係数番号 1、19 は比較的分離できている。ただ、番号 1 は直流成分なので、手の大きさや距離によって変化が出てくる可能性は否定できない。3 番や 6 番は「チョキ」と「グー」に関しては重なってしまっているが、「チョキ」や「グー」と「パー」の間では比較的大きい距離があり、「グー」と「パー」の 2 要素で分離したい場合などでは安定した結果が出ると考えられる。

また、1 つの特徴では 2 つしか分離できなくても、2 つ以上の特徴を組み合わせると分離できる場合場合がある。

このように特定の係数番号をピックアップして比較すれば、回転に対応したパターン認識が出来る可能性がある。ただし、「グー」、「チョキ」、「パー」の 3 種のジェスチャーに対して的確に分離できると考えるには、アルゴリズムやシステムに対して改良や見直しの余地があると思われる。

3.3 手の傾きに対する DCT 係数の変化

傾きに対するシミュレーションにおいても、回転シミュレーションと同様に式(3.1)、(3.2)、(3.3)を用いて DCT 係数を計算する。手の傾きに対するシミュレーションでは、図 3.6 の画像計 36 枚を使用する。サイズは 8×8 であり、得られる DCT 係数は 64 個である。3.2 の回転シミュレーションと同様に、各画像の DCT 係数番号の振幅変化を評価していく。グラフ内における系列番号は、図 36 内の画像番号を示している。

まず、「グー」に関する DCT 係数の振幅をグラフにしたものを図 3.11 に示す。

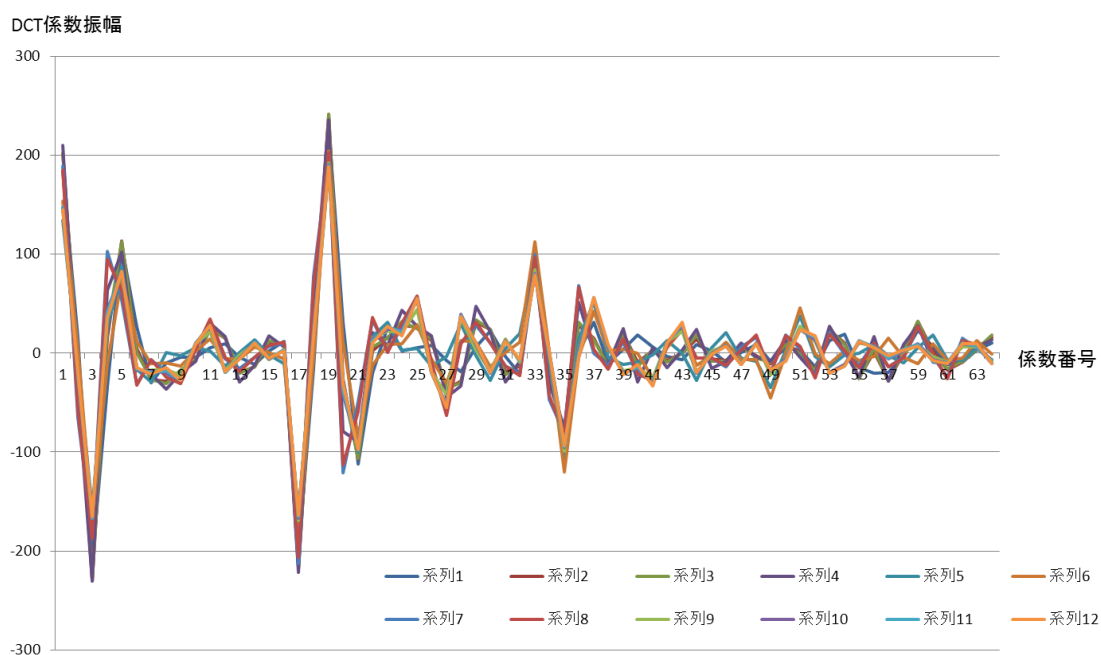


図 3.11 「グー」の手の傾きに対する DCT 係数の振幅

傾きによって形状が異なるので、3.7.1 の回転に関する振幅のグラフと比べると、やはり振幅差が大きい。また、直流成分である係数番号 1 番も形状によって輝度値の高い部分の面積が変化するので振れ幅が大きい。ただ、番号 3、17、19、33 等は比較的振れ幅は小さい。係数が大きく成る程、高周波領域を表現した値であり、値は全体的に小さくなっていく。よって、比較的小さい係数番号の変化量の方が特徴量の判別には向いていると言える。

次に「チョキ」に関する DCT 係数の振幅をグラフにしたものを図 3.12 に示す。

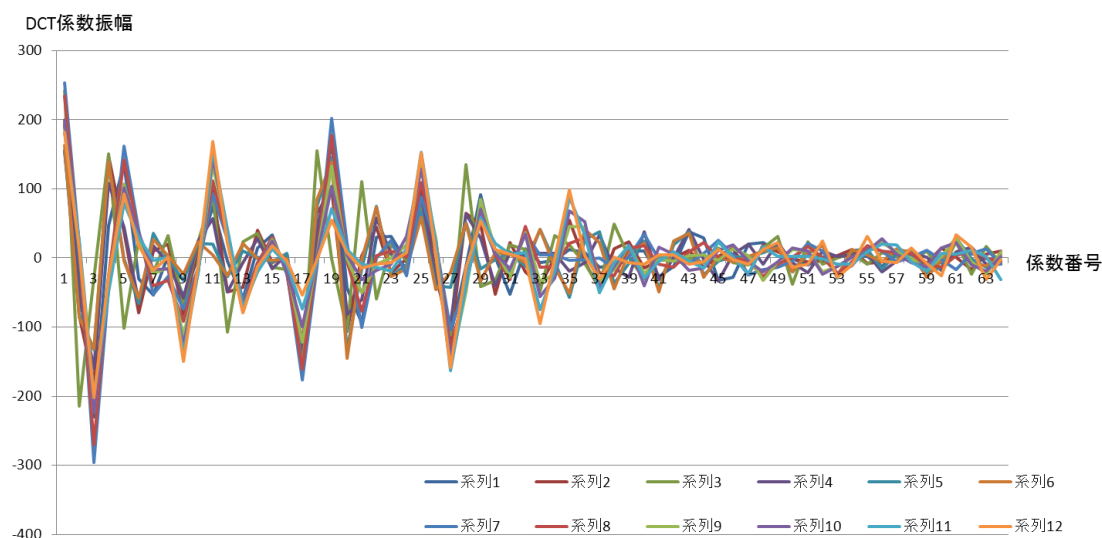


図 3.12 「チョキ」の手の傾きに対する DCT 係数の振幅

「チョキ」になると、低い係数番号から大きく振幅差が出ており、判別は難しいと思われる。特に縦方向に深く傾けた場合(系列 10～12)、形状が他の画像と大きく異なってしまう、視覚的に考えても「グー」に近くなってしまうので、こういった傾きに対しての判別難しいと言わざるを得ない。もし、これに対応する様にするには、傾けた場合のモデルデータをこと前に用意し、ラベル付けするといったこと以外では難しいだろう。

最後に「パー」に関する DCT 係数の振幅をグラフにしたものを図 3.13 に示す。

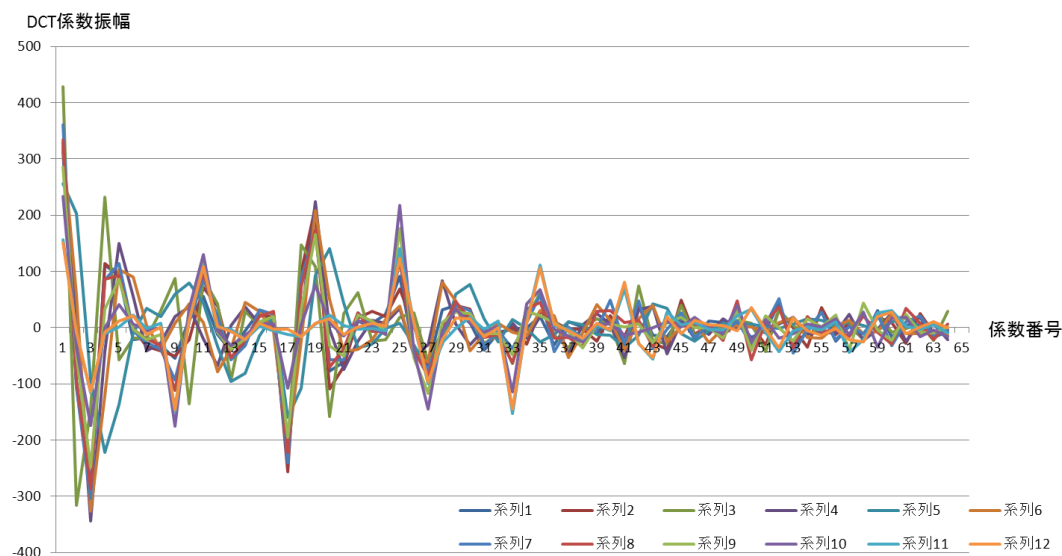


図 3.13 「パー」の手の傾きに対する DCT 係数の振幅

「パー」も「チョキ」と同様に非常に振れ幅が大きい。よって、傾きの変化が大きい場合、DCT 係数を用いた判別法では対応は難しいと思われる。

あくまでこの手法は画像を直接 DCT 変換してその係数を比較しているに過ぎないので、回転などの変化に弱いということがアルゴリズム上の原理的な問題として考えられる。この問題に対応するには、画像の信号を他の特徴量に一度変換して入力する必要があると思われ、そういった変換や処理は、次章の「フーリエ記述子を用いたパターン認識」で行っている。

DCT 係数の用いた判別法は、モデル画像に比較的近い角度、傾きの際では有効な可能性もあるが、動く対象など常に同じ条件を保つのが難しいような対象では、DCT 係数の安定性を保つのは非常に難しいと予想される。

この手法を用いた判別システムを確立させようとする場合、考えられる条件のモデル画像に対する DCT 係数の特徴ベクトルモデル数を用意し、それぞれの DCT 係数をひとつひとつに対してラベル付けを行いモデルデータベースの拡張対応するのが 1 つの解決方法であると考えられる。

第 4 章

フーリエ記述子を用いた回転に対するパターン認識シミュレーション

4.1 使用する低解像度熱赤外面像

本章のシミュレーションで使用する画像は、3 章で用いた画像と同じものである。これらは「グー」「チョキ」「パー」の三種類のジェスチャーの温度情報を熱赤外センサで取得したものである。ここでは、3 章における DCT 係数による評価における問題とされていた回転がある場合のジェスチャー認識に焦点を当ててシミュレーションを行う。

まず基本となる画像として図 4.1 の 3 種類の画像を用意する。その後、不必要部分のクリッピングを行い、図 4.2 の 3 枚の画像に変換する。温度情報のクリッピングでは閾値レベルを 90 と設定した。

これらの画像を回転角度 0° とし、反時計回りで 10° 刻みで 90° まで回転した画像 9 枚ずつを画像の回転によって作成する。よって、基本画像 3 枚 + 回転画像 27 枚の計 30 枚の画像を用意する(図 4.3)。また、回転によって手の長軸部分の情報が失われるのを防ぐために、横 3 画素、縦 4 画素の余白(輝度値 0 空間)を付加した上で回転処理を行う。

次に、これら 51×51 画素の 30 枚の画像に対して縮小処理を行い、 16×16 画素と 8×8 画素の 2 種類のサイズの低解像度熱赤外面像を用意する。縮小処理には線形補間処理アルゴリズムを用いて行う。図 4.4 は 16×16 画素の低解像度熱赤外面像群であり、図 4.5 は 8×8 画素の熱赤外面像群である。



図 4.1 「グー」「チョキ」「パー」の基本画像

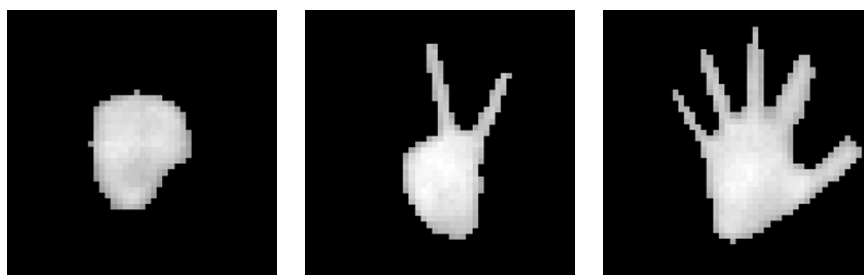


図 4.2 基本画像の抽出画像

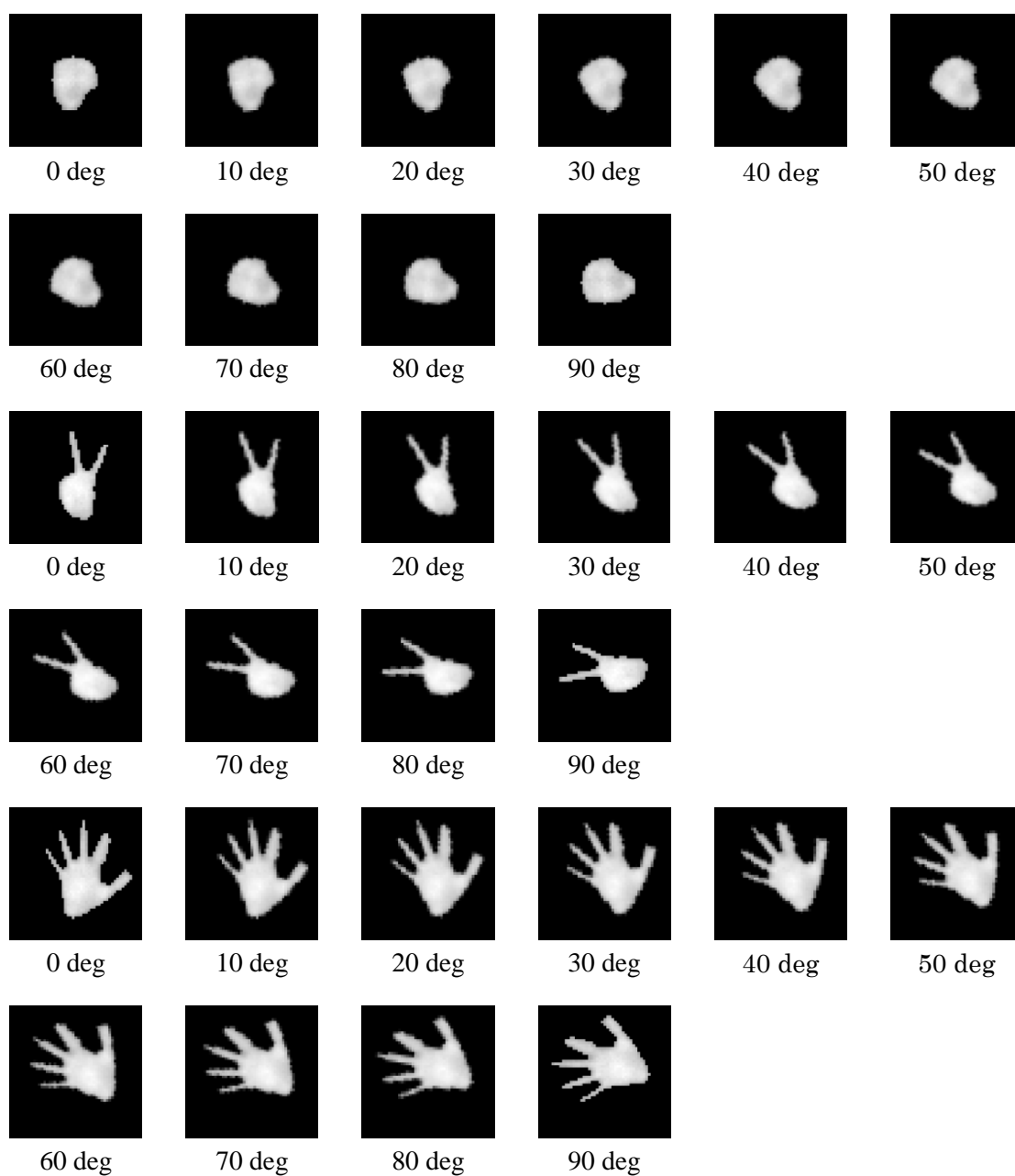


図 4.3 「グー」「チョキ」「パー」の回転画像 30 枚(51×51 画素)

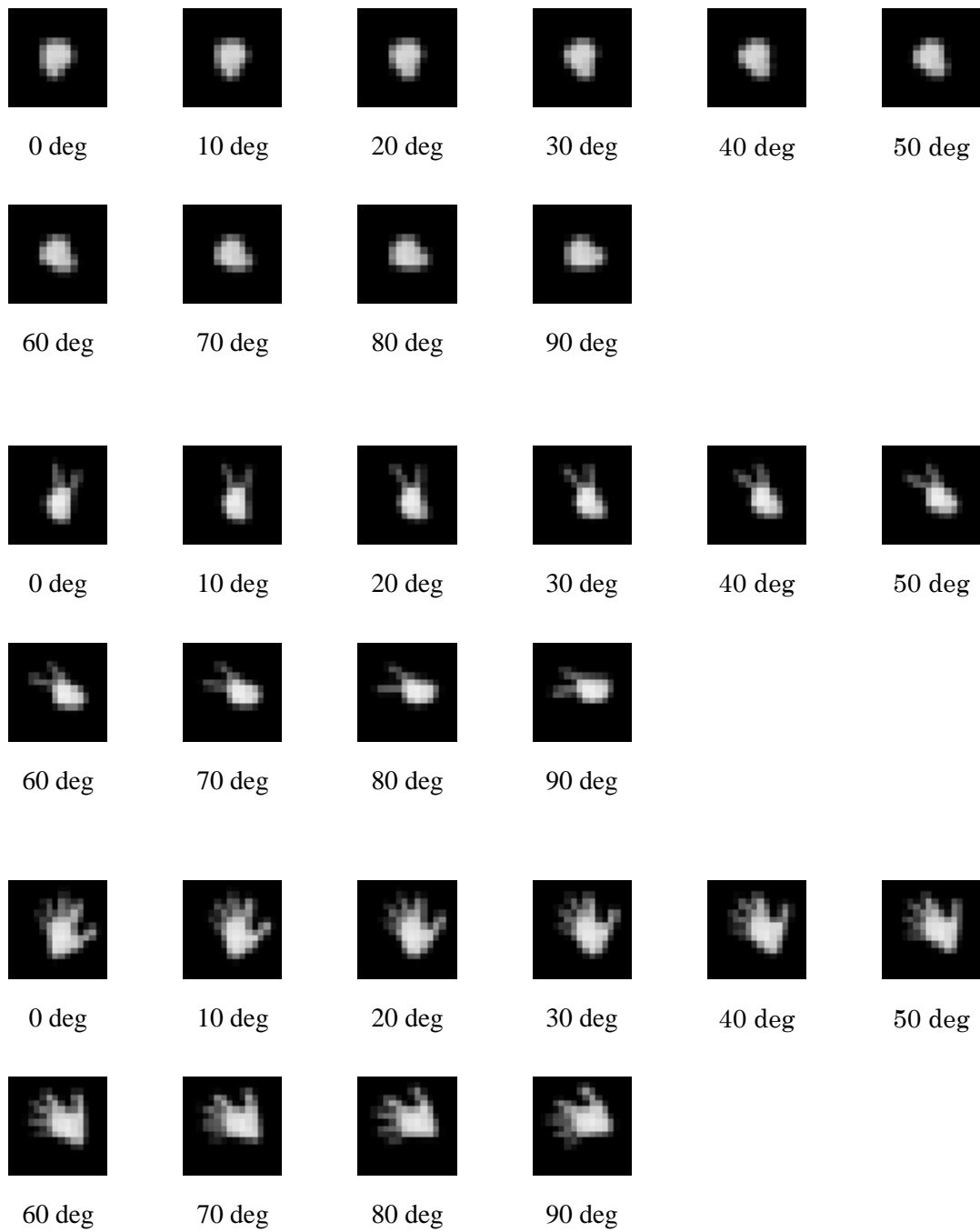


図 4.4 16×16 画素の低解像度画像

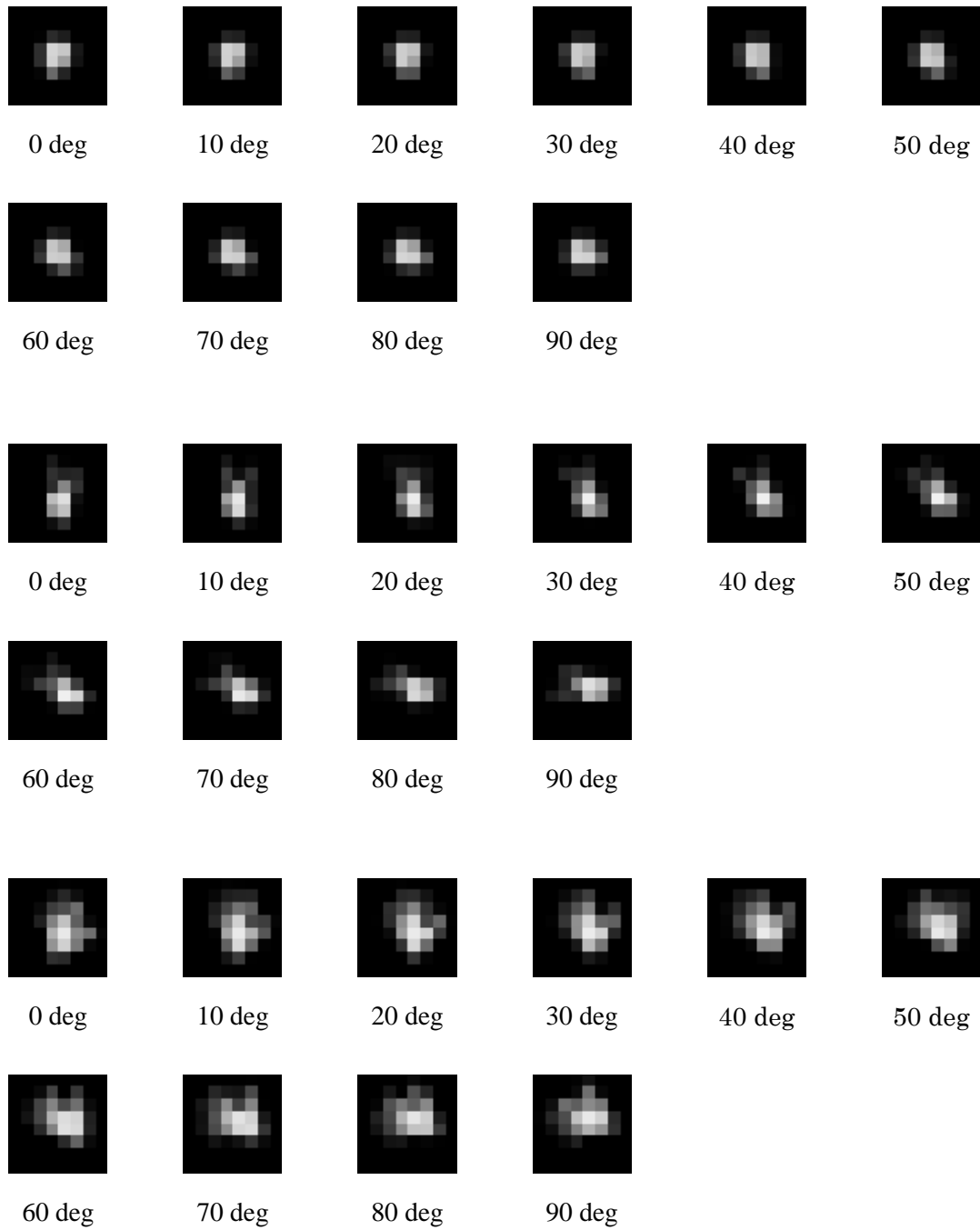


図 4.5 8×8 画素の低解像度画像

4.2 輪郭抽出と代表点の抽出シミュレーション

フーリエ記述子を用いたパターン認識を行う上で、離散フーリエ変換の式(2.9)や(2.10)に入力する周期関数を与えるために、画像の輪郭情報を用いる。また、それらの特徴量を保存しつつデータ処理の簡便性を図るため、輪郭情報のある間隔でサンプリングした代表点を抽出する必要がある。

まず、輪郭線の抽出を行う。4.1 節で作成した低解像度画像(図 4.4、図 4.5)に対して低解像度画像のまま輪郭を抽出するのは困難である。この問題を解決するため、一度低解像度画像から 128×128 画素の高解像度画像へと拡大処理を行う。この処理は補間によって高解像度画像を作成しているので、情報量は増えていない。ここで、拡大処理アルゴリズム(文献[10]))として、**Bi-Cubic** 補間法と **Bi-Linear** 補間法のどちらを採用するか検討を行う。図 4.6 は 16×16 ピクセルの角度 0° 条件の「チョキ」に対して補間拡大処理を行ったものである。右図の **Bi-Linear** による画像に比べて左図の **Bi-Cubic** を用いて補間処理を行ったものの方が滑らかな輝度勾配になっているのが解る。これは、**Bi-Linear** 手法を用いた場合、元の画素サイズが大きいと補間の結果に影響が現れやすいことが考えられる。そのため、画像の所々で「角」と取れる輝度勾配が見受けられる。**Bi-Cubic** を使用した場合、アルゴリズムの特性上、輝度値がオーバーシュートする可能性も考えられるが、輪郭抽出の際はできるだけ滑らかな輝度勾配の方が好ましいので、本研究では、拡大処理アルゴリズムとして **Bi-Cubic** 補間法を使用することとする。図 4.7、図 4.8 は、**Bi-Cubic** 拡大処理によって作成した画像群である。



図 4.6 拡大画像(128×128 画素) 一左 : **Bi-Cubic** 補間手法、右 : **Bi-Linear** 補間手法



図 4.7 16×16 画素低解像度画像の高解像度補間画像(128×128)

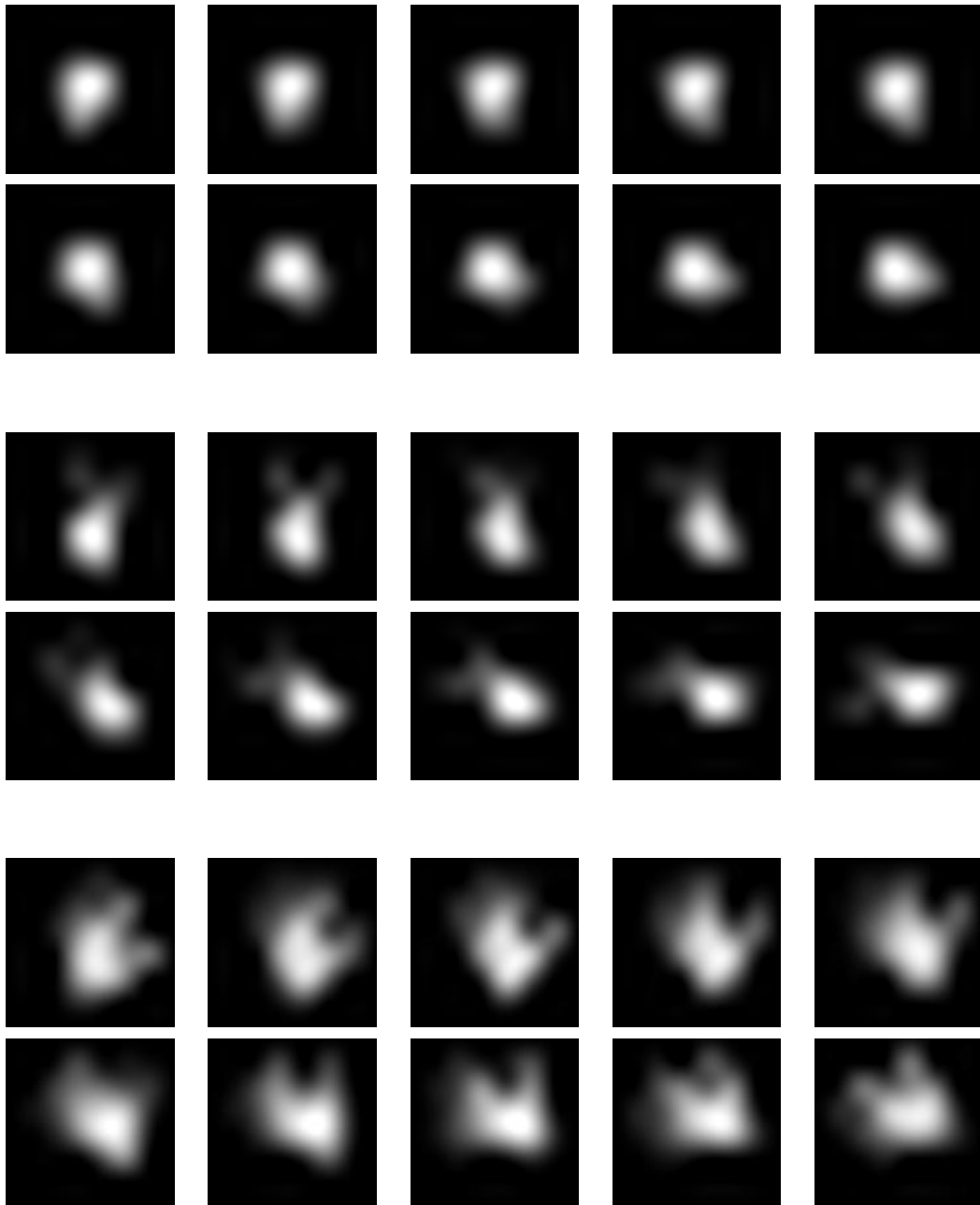


図 4.8 8×8 画素低解像度画像の高解像度補間画像(128×128)

次に拡大処理を行った画像に対して輪郭抽出を行う。輪郭抽出方法は、第 2 章の 2.4.2

で述べたものを用いる。輪郭抽出の際に用いる閾値LKは、視覚的に判断して最も元画像の形状に近く抽出された値とした。具体的には、 16×16 画素の場合 $LK_{16}=50$ 、 8×8 画素の場合 $LK_8=40$ である。輪郭抽出によって得られた輪郭画像を図 4.9、図 4.10 で示す。

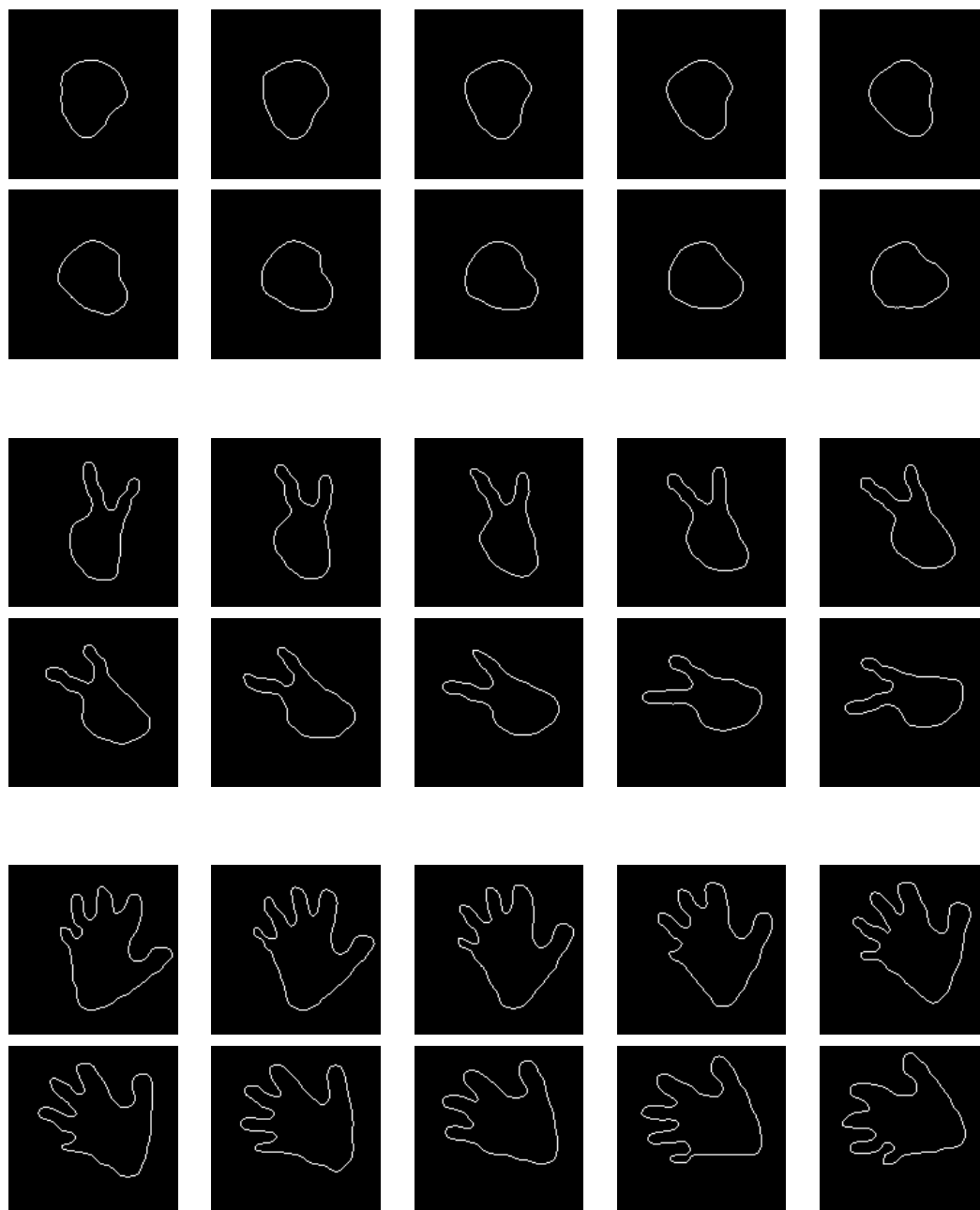


図 4.9 16×16 画素低解像度画像を元にした輪郭抽出画像(128×128)

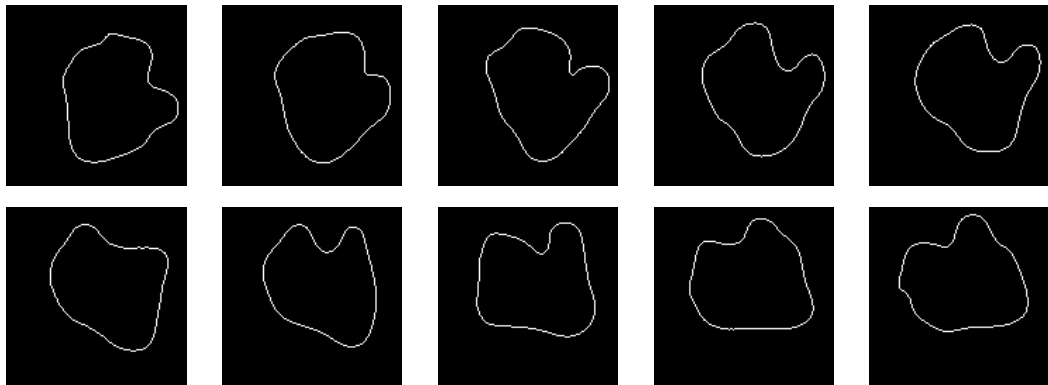
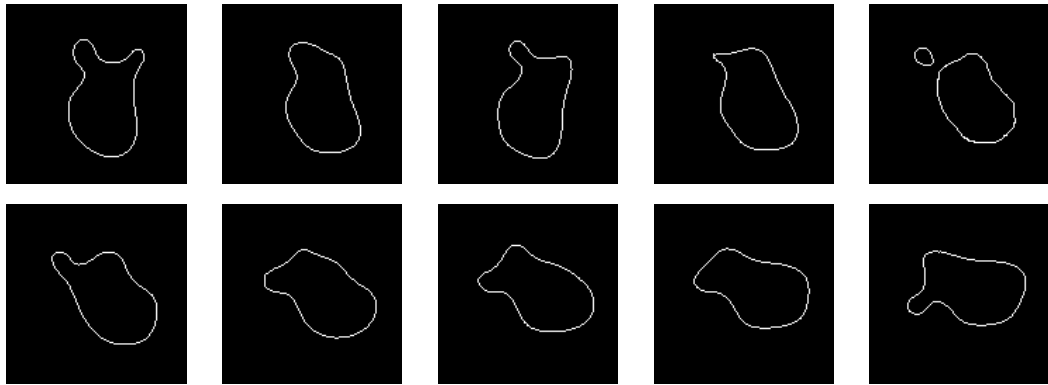
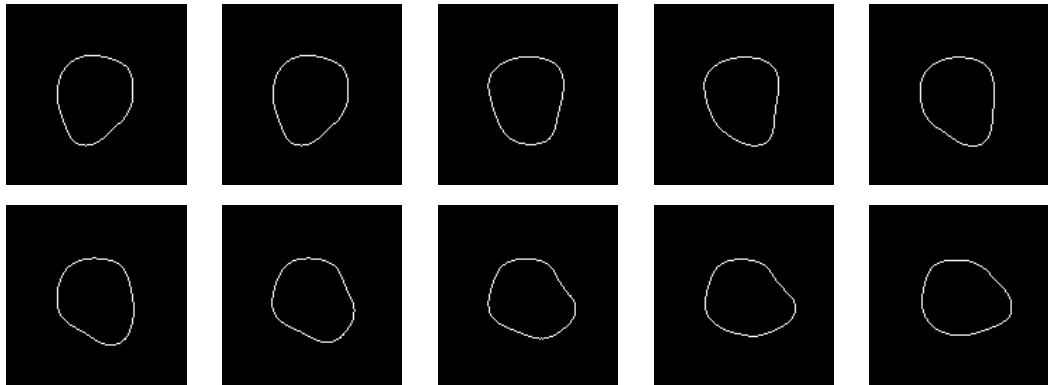


図 4.10 8×8 画素低解像度画像を元にした輪郭抽出画像(128×128)

輪郭抽出した画像に対して代表点をサンプリングし、入力データ数の削減を行う。その際、データ削減によって特徴量が失われないように注意する。サンプリング方法としては 2.4.3 の(2)の手法を用いる。この手法を選択する理由として、(1) 等間隔で代表点のサンプリングを行いたい、(2)始点間と終端間を等間隔でとりたい、(3)異なる画像間においてはサンプリング距離の一致は必要ない、という 3 つの理由からである。

この手法を画像処理に実際に適用にするにあたっていくつかの問題がある。まず一つ目は、輪郭の全長を測定する方法、そして二つ目はサンプリング点の決定条件の設定である。

まず一つ目の問題を解決するために、輪郭の全長の測定の際、輪郭の画素数を始点から終端までカウントすることでそのカウント数を全長することとする。実際、上下左右に関しては距離 1、斜めの距離は約 1.41 であり、4 近傍か 8 近傍かによって長さが異なる。本研究ではシミュレーションの簡便のため、画素間の距離は全て 1 と近似する。

そして、輪郭を構成する画素の全カウント数を L とし、入力点数を N とし(2.12)に適用する。

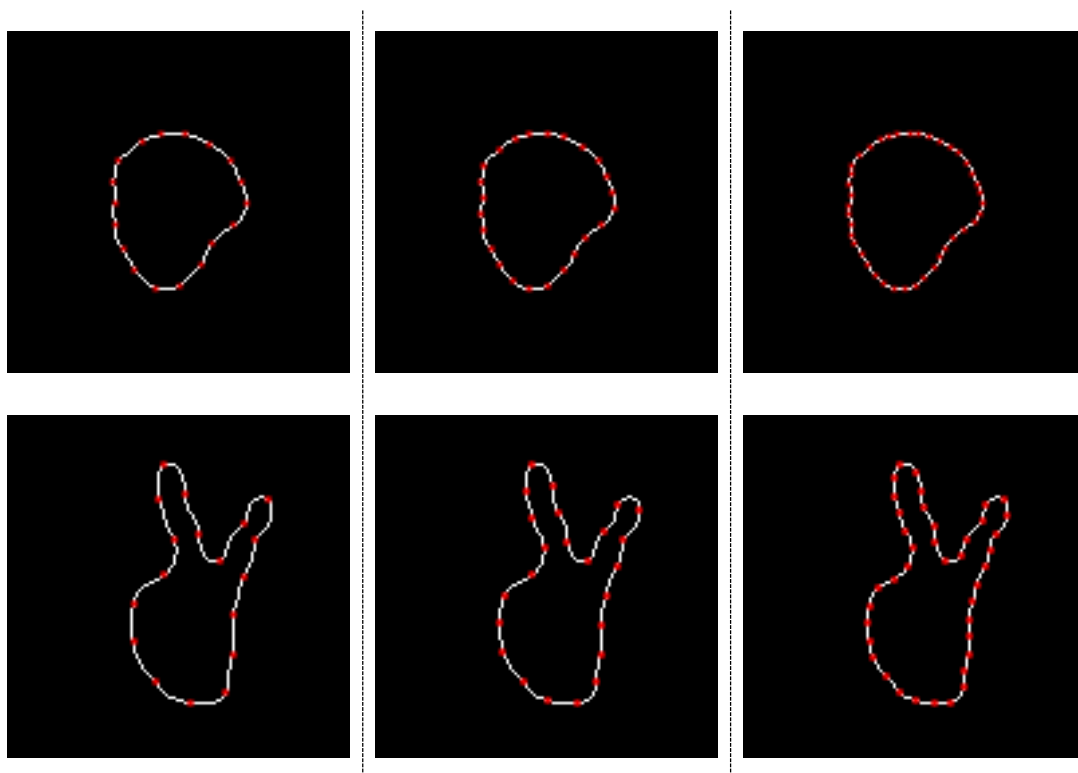
次に 2 つ目の問題への対応を行う。サンプリング間隔 l が常に整数とは限らないため、以下の式を用いて代表点座標の決定を行う。

$$l_0 = \frac{L_0}{N}, \quad S = \sum_{i=0}^{N-1} l_0 i, \quad \bar{S} = \text{round}(S) \quad (4.1)$$

この式において、 L_0 は全長のカウント数、 N はサンプリング点数、 l_0 はサンプリング間隔、 S は始点からの小数点付きカウント数、 \bar{S} は S の値を四捨五入し整数化したカウント数である。サンプリングにおいて、始点からの輪郭画素のカウント数 \bar{S} 個目に該当するものを代表点とする。図 4.11 は、図 4.9、図 4.10 の画像の一部と代表点画像とのレイヤーの重ね合わせを用いて、適切にサンプリングできているかを示す。また、図 4.12、図 4.13 は 0° 条件時の「ゲー」「チョキ」「パー」画像に対して、サンプリング点数 $N=18, 24, 30, 36$ に変化させて抽出した代表点の座標を画像化したものである。

本研究では、サンプリング点数による判別結果の変化を見るために、サンプリング点数 N を 18 から 2 刻みで 40 まで変えて認識シミュレーションを行う。加えて回転による認識精度の変化を見るために 10 通り、回転による変化を見るために 10 通り、解像度による変化を見るために 16×16 画素と 8×8 画素の 2 種類の画像を使用するため、計 $2 \times$

$10 \times 12 = 240$ 通りの場合についてシミュレーションを行う。



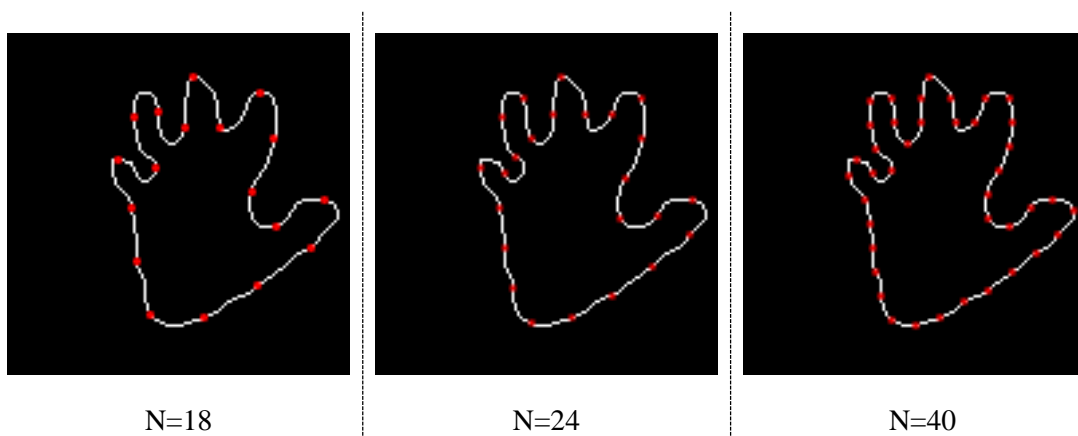


図 4.11 輪郭からサンプリングした代表点

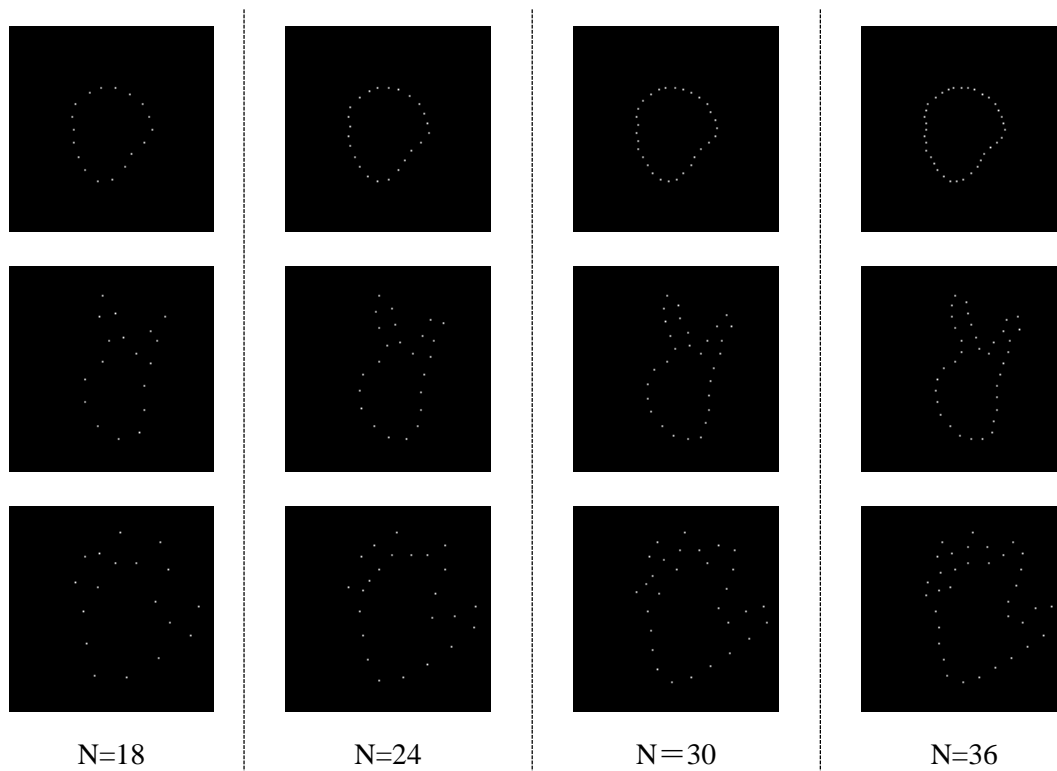
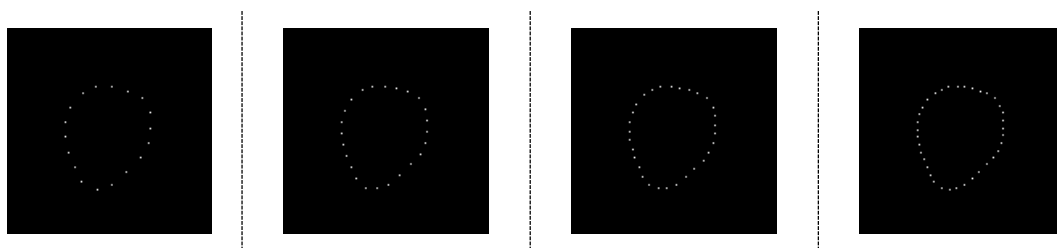


図 4.12 サンプリング点数 N による代表点の座標(16×16)



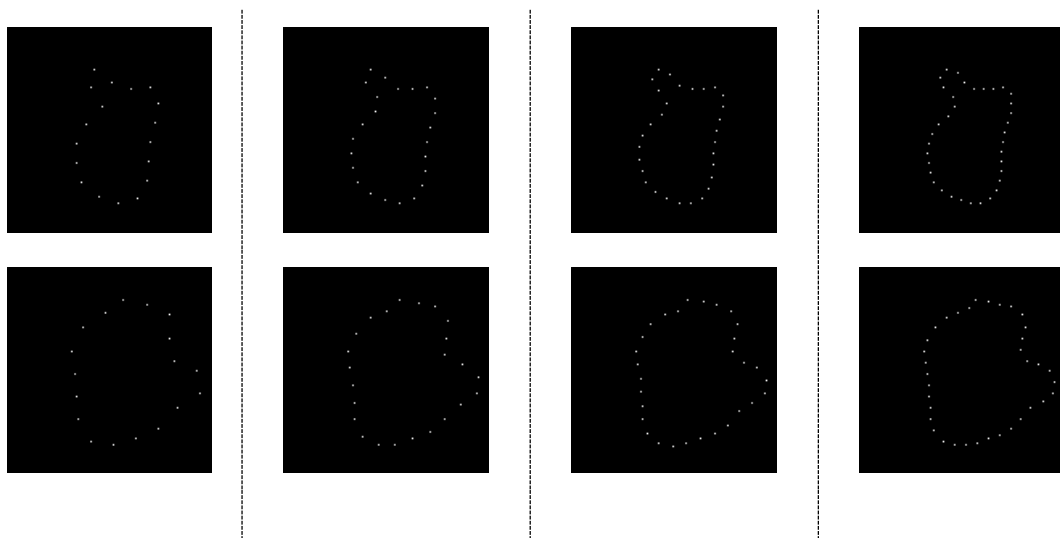


図 4.13 サンプルング点数 N による代表点の座標(8×8)

4.3 画像の回転と代表点数を変化させた場合の認識シミュレーション

図 4.9、図 4.10 で示される各 30 枚の回転した低解像度画像に対して、入力点数を変化させながら認識シミュレーションを行う。また、認識アルゴリズムには Z 型フーリエ記述子を用いたパターン認識手法と P 型フーリエ記述子を用いたパターン認識手法の 2 種類を使用する。この試験は、回転による認識精度の検証、サンプルング点数による認識精度の検証、そして認識手法の評価を行うためのものである。

シミュレーションの流れを図 4.14 に示す。まず、代表点数を設定する。代表点数はモデルデータ、認識処理ともに共通の数とする。次に、「グー」「チョキ」「パー」それぞれの 0° 画像をモデル画像とし、図 4.14 における「①<モデルデータ作成>」の流れで処理を行う。結果として出たフーリエ記述子を特徴ベクトルとし、データベースに登録を行う。その際、それぞれの結果に「グー」、「チョキ」、「パー」のラベル付けを行う。次に、認識処理を図 4.14 の「②<認識処理>」の流れで行う。この処理の結果として出力された特徴ベクトルと事前にデータベースに登録しておいた特徴ベクトルとの間の評価を行い、認識結果を出力する。

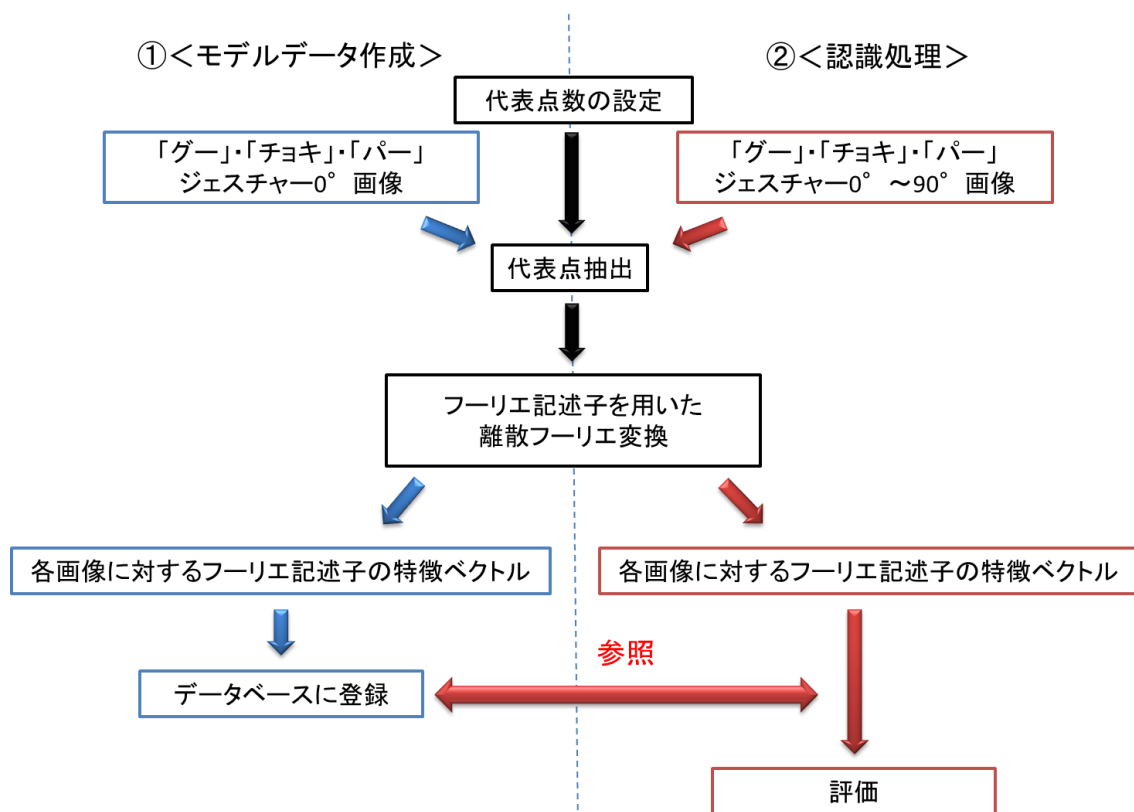


図 4.14 認識シミュレーションの流れ

4.3.1 Z 型フーリエ記述子パターン認識を用いたシミュレーション

まず、第 2 章 2.4.1 (ii) で説明した偏角関数表現である Z 型フーリエ記述子を用いたパターン認識シミュレーションを行う。Z 型フーリエ記述子で使用する周期関数は、

$$\varphi_i = \theta_i - \theta_0 - \frac{2\pi}{N}i \quad (i = 0, 1, 2, \dots, N-1) \quad (2.8)$$

である。シミュレーションでは、 $N=18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40$ の 12 種類の点数で代表点抽出を行った。また、評価する際、低周波成分の $k=1 \sim k=6$ までのフーリエ記述子を用いた。

例として、図 4.15 は 16×16 画素の「ゲー」「チョキ」「パー」 0° 基本画像のフーリエ係数である。図 4.15 のグラフは出力された係数の実数部と虚数部の絶対値である。

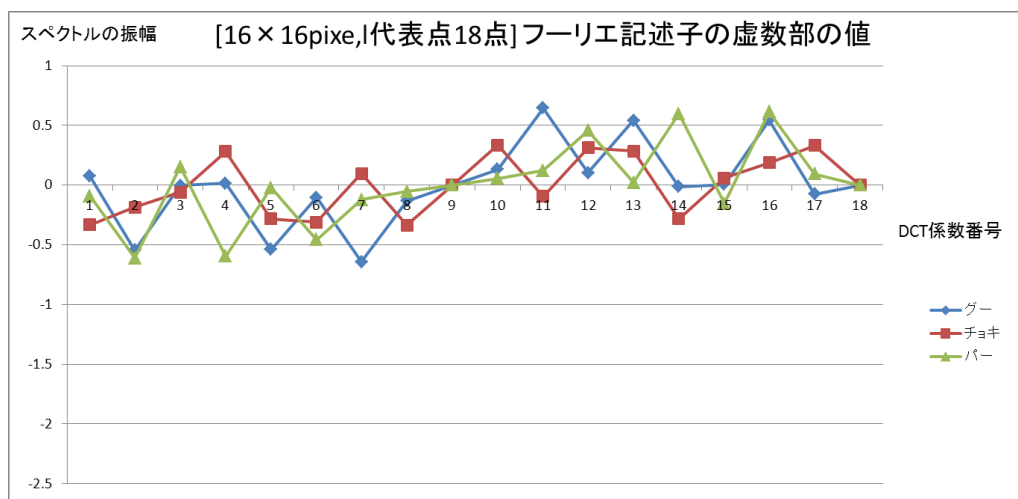
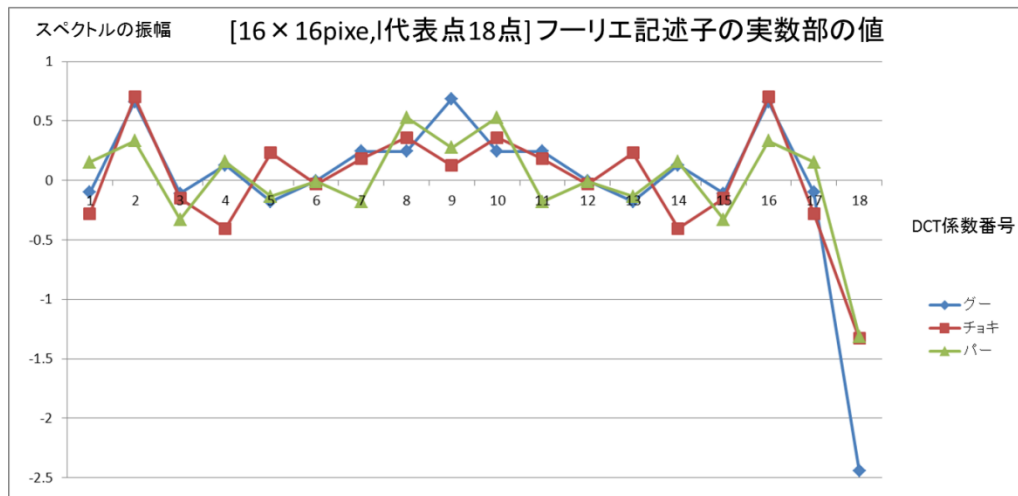


図 4.15 フーリエ記述子の振幅(16×16、代表点 18 点)

また、図 4.16 は 8×8 画素の「グー」「チョキ」「パー」0° 基本画像のフーリエ係数である。図 4.15 と同様に、上のグラフは実数部、下のグラフは虚数部を示す。

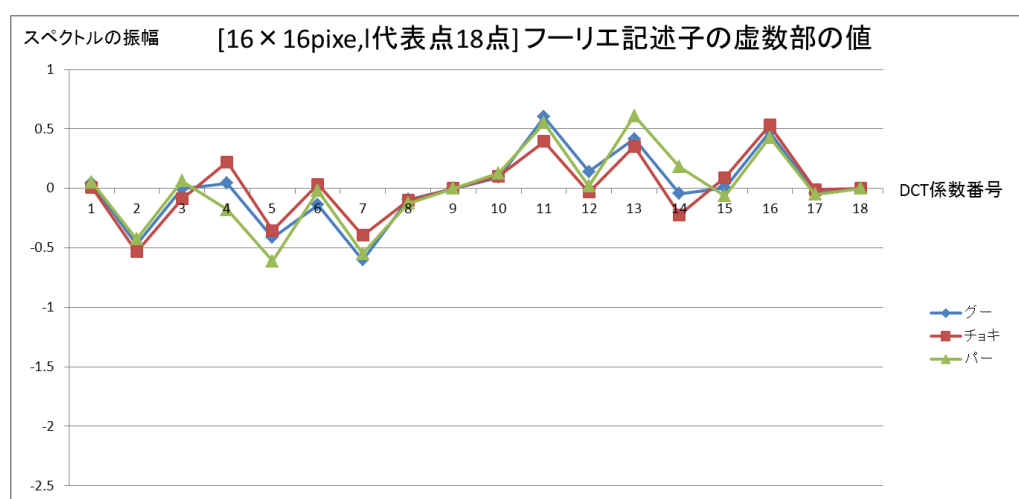
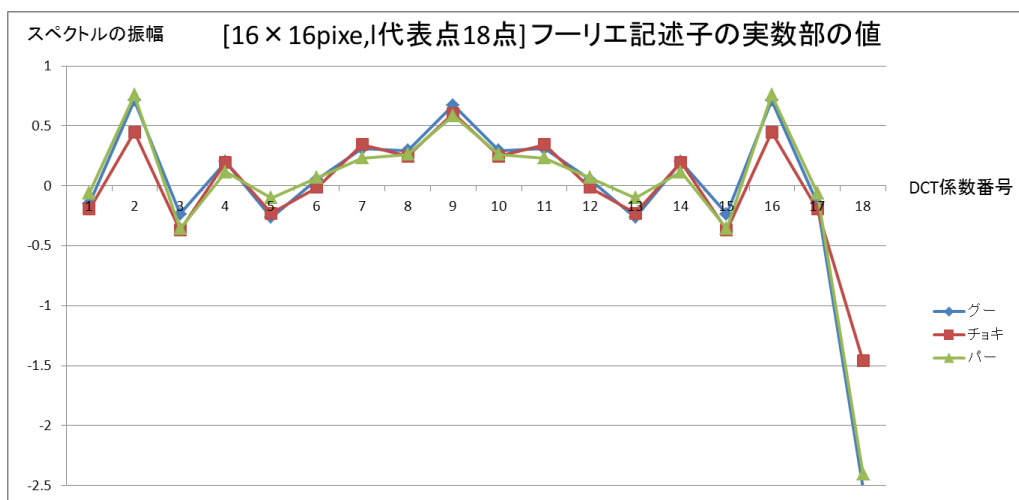


図 4.16 フーリエ記述子の振幅(8×8、代表点 18 点)

次に認識シミュレーションを行うため、モデルデータの特徴ベクトルと認識対象画像の特徴ベクトルとの評価を行う。評価には、第 2 章における 2.4.4 の評価式を用い、モデル特徴ベクトルと認識画像の特徴ベクトルにおける各係数のユークリッド距離が最も小さいものを答えとして出力する。

本項の表 4.1、表 4.2 で示すのは、その結果の正答率である。

ゲー				チョキ				パー			
N	正解	不正解	正答率	N	正解	不正解	正答率	N	正解	不正解	正答率
18	10	0	100	18	3	7	30	18	5	5	50
20	10	0	100	20	4	6	40	20	5	5	50
22	10	0	100	22	4	6	40	22	5	5	50
24	10	0	100	24	4	6	40	24	5	5	50
26	10	0	100	26	3	7	30	26	5	5	50
28	7	3	70	28	3	7	30	28	10	0	100
30	10	0	100	30	4	6	40	30	4	6	40
32	7	3	70	32	3	7	30	32	10	0	100
34	10	0	100	34	4	6	40	34	4	6	40
36	10	0	100	36	3	7	30	36	6	4	60
38	10	0	100	38	3	7	30	38	3	7	30
40	10	0	100	40	3	7	30	40	6	4	60

N	総合正答率
18	60.00
20	63.33
22	63.33
24	63.33
26	60.00
28	66.67
30	60.00
32	66.67
34	60.00
36	63.33
38	53.33
40	63.33
平均	61.94

表 4.1 Z 型フーリエ記述子正答表：16×16 画素における代表点*N*ごとの正当結果

ゲー				チョキ				パー			
N	正解	不正解	正答率	N	正解	不正解	正答率	N	正解	不正解	正答率
18	8	2	80	18	3	7	30	18	2	8	20
20	3	7	30	20	3	7	30	20	6	4	60
22	2	8	20	22	6	4	60	22	2	8	20
24	6	4	60	24	3	7	30	24	3	7	30
26	5	5	50	26	3	7	30	26	3	7	30
28	9	1	90	28	3	7	30	28	4	6	40
30	7	3	70	30	3	7	30	30	2	8	20
32	3	7	30	32	6	4	60	32	6	4	60
34	5	5	50	34	4	6	40	34	2	8	20
36	3	7	30	36	4	6	40	36	1	9	10
38	4	6	40	38	6	4	60	38	2	8	20
40	6	4	60	40	6	4	60	40	2	8	20

N	総合正答率
18	43.33
20	40.00
22	33.33
24	40.00
26	36.67
28	53.33
30	40.00
32	50.00
34	36.67
36	26.67
38	40.00
40	46.67
平均	40.56

表 4.2 Z 型フーリエ記述子正答表：8×8 画素における代表点*N*ごとの正当結果

表 4.1、表 4.2 から判るように、正答率は決して高いとは言えない。8×8 画素に比べて若干 16×16 画素の正答率が高く、最高正答率 66%であった。しかし、決して実際の運用に堪える正答率とは言えない。比較的「グー」の正答率が高いが、特に「チョキ」に関する正答率が低い。

16×16 画素に関して表 4.3.1 と表 4.3.2、8×8 画素に関して表 4.4.1 と表 4.4.2 において、最も正答率の高い N と最も正答率の低い N に関する各ジェスチャーのユークリッド距離、回答カテゴリを示す。

角度	ラベル	距離	答え
0°	グー	0	グー
	チョキ	0.364679	
	パー	0.394935	
10°	グー	0.266586	グー
	チョキ	0.536003	
	パー	0.41392	
20°	グー	0.645045	グー
	チョキ	0.857568	
	パー	0.667589	
30°	グー	0.365102	グー
	チョキ	0.638436	
	パー	0.462814	
40°	グー	0.399146	グー
	チョキ	0.618659	
	パー	0.425355	
50°	グー	0.369929	グー
	チョキ	0.594264	
	パー	0.384797	
60°	グー	0.566591	
	チョキ	0.770299	
	パー	0.54755	パー
70°	グー	0.55971	
	チョキ	0.761719	
	パー	0.548702	パー
80°	グー	0.849019	
	チョキ	1.029428	
	パー	0.836046	パー
90°	グー	0.107564	グー
	チョキ	0.341983	
	パー	0.405243	

角度	ラベル	距離	答え
0°	グー	0.364679	
	チョキ	0	チョキ
	パー	0.33679	
10°	グー	0.650758	
	チョキ	0.35588	チョキ
	パー	0.661355	
20°	グー	0.579959	
	チョキ	0.314534	チョキ
	パー	0.612557	
30°	グー	1.230617	
	チョキ	1.249521	
	パー	1.056864	パー
40°	グー	1.179703	
	チョキ	1.220202	
	パー	1.033714	パー
50°	グー	1.137931	
	チョキ	1.189614	
	パー	0.992846	パー
60°	グー	1.159683	
	チョキ	1.221519	
	パー	1.020389	パー
70°	グー	0.811541	
	チョキ	0.893894	
	パー	0.678045	パー
80°	グー	0.898344	
	チョキ	0.960281	
	パー	0.770758	パー
90°	グー	1.223834	
	チョキ	1.107441	チョキ
	パー	1.158886	

角度	ラベル	距離	答え
0°	グー	0.394935	
	チョキ	0.33679	
	パー	0	パー
10°	グー	0.489187	
	チョキ	0.412418	
	パー	0.157875	パー
20°	グー	1.010321	
	チョキ	1.089547	
	パー	0.975912	パー
30°	グー	0.632848	
	チョキ	0.686105	
	パー	0.58092	パー
40°	グー	0.907637	
	チョキ	0.993	
	パー	0.864981	パー
50°	グー	0.858852	
	チョキ	0.957283	
	パー	0.844693	パー
60°	グー	1.078178	
	チョキ	1.161118	
	パー	1.059887	パー
70°	グー	1.786687	
	チョキ	1.651415	
	パー	1.529141	パー
80°	グー	1.17547	
	チョキ	0.981941	
	パー	0.901072	パー
90°	グー	0.872638	
	チョキ	0.934499	
	パー	0.747948	パー

表 4.3.1 16×16 画素において正答率が最も高かった N=32

各角度にたいするユークリッド距離と回答カテゴリ(左：グー、中央：チョキ、右：パー)

角度	ラベル	距離	答え
0°	グー	0	グー
	チョキ	0.798864	
	パー	0.695845	
10°	グー	0.302214	グー
	チョキ	0.965906	
	パー	0.821757	
20°	グー	0.550621	グー
	チョキ	1.157418	
	パー	1.017229	
30°	グー	0.325003	グー
	チョキ	0.96128	
	パー	0.807779	
40°	グー	0.355338	グー
	チョキ	0.96022	
	パー	0.806817	
50°	グー	0.329143	グー
	チョキ	0.919573	
	パー	0.748472	
60°	グー	0.484837	グー
	チョキ	1.117788	
	パー	0.974033	
70°	グー	0.462956	グー
	チョキ	1.109212	
	パー	0.982816	
80°	グー	0.69937	グー
	チョキ	1.348645	
	パー	1.23497	
90°	グー	0.427344	グー
	チョキ	1.10355	
	パー	1.001964	

角度	ラベル	距離	答え
0°	グー	0.798864	
	チョキ	0	チョキ
	パー	0.350949	
10°	グー	0.92261	
	チョキ	0.297256	チョキ
	パー	0.549745	
20°	グー	0.893071	
	チョキ	0.380834	チョキ
	パー	0.578097	
30°	グー	1.14138	グー
	チョキ	1.413553	
	パー	1.301798	
40°	グー	1.053307	グー
	チョキ	1.351867	
	パー	1.230977	
50°	グー	0.975522	グー
	チョキ	1.339924	
	パー	1.200619	
60°	グー	0.935335	グー
	チョキ	1.325326	
	パー	1.186322	
70°	グー	0.582776	グー
	チョキ	1.136701	
	パー	0.990377	
80°	グー	0.521903	グー
	チョキ	1.014864	
	パー	0.900786	
90°	グー	0.975728	グー
	チョキ	1.033483	
	パー	1.117531	

角度	ラベル	距離	答え
0°	グー	0.695845	
	チョキ	0.350949	
	パー	0	パー
10°	グー	0.72673	
	チョキ	0.412769	
	パー	0.101989	パー
20°	グー	0.713964	グー
	チョキ	1.159479	
	パー	1.104773	
30°	グー	0.49584	グー
	チョキ	0.986723	
	パー	0.943775	
40°	グー	0.773849	グー
	チョキ	1.256858	
	パー	1.20488	パー
50°	グー	0.492133	グー
	チョキ	1.047451	
	パー	0.990193	
60°	グー	0.703241	グー
	チョキ	1.226543	
	パー	1.160136	
70°	グー	1.726536	
	チョキ	1.568528	
	パー	1.510644	パー
80°	グー	0.960947	
	チョキ	0.656963	
	パー	0.72598	パー
90°	グー	0.880504	
	チョキ	1.093184	
	パー	0.911133	パー

表 4.3.2 16×16 画素において正答率が最も低かった N=38

各角度に対するユークリッド距離と回答カテゴリ(左：グー、中央：チョキ、右：パー)

角度	ラベル	距離	答え
0°	グー	0	グー
	チョキ	0.378785	
	パー	0.407568	
10°	グー	0.372394	
	チョキ	0.278183	
	パー	0.228301	パー
20°	グー	0.11595	グー
	チョキ	0.402165	
	パー	0.436555	
30°	グー	0.22543	グー
	チョキ	0.433803	
	パー	0.472753	
40°	グー	0.423511	グー
	チョキ	0.714687	
	パー	0.790386	
50°	グー	0.423593	グー
	チョキ	0.714973	
	パー	0.769486	
60°	グー	0.41474	グー
	チョキ	0.717296	
	パー	0.75828	
70°	グー	0.372162	グー
	チョキ	0.699991	
	パー	0.703127	
80°	グー	0.264086	グー
	チョキ	0.511491	
	パー	0.437925	
90°	グー	0.482853	
	チョキ	0.410756	
	パー	0.1819	パー

角度	ラベル	距離	答え
0°	グー	0.378785	
	チョキ	0	チョキ
	パー	0.346923	
10°	グー	0.666158	
	チョキ	0.401025	チョキ
	パー	0.523552	
20°	グー	0.304858	
	チョキ	0.299037	チョキ
	パー	0.410151	
30°	グー	0.717681	グー
	チョキ	1.002749	
	パー	1.065292	
40°			
50°	グー	0.892948	
	チョキ	0.643082	
	パー	0.633242	パー
60°	グー	0.321686	グー
	チョキ	0.689913	
	パー	0.644752	
70°	グー	0.261529	グー
	チョキ	0.636625	
	パー	0.627643	
80°	グー	0.594599	
	チョキ	0.566892	
	パー	0.312283	パー
90°	グー	0.44474	グー
	チョキ	0.55713	
	パー	0.510986	

角度	ラベル	距離	答え
0°	グー	0.407568	
	チョキ	0.346923	
	パー	0	パー
10°	グー	0.387975	グー
	チョキ	0.728218	
	パー	0.748341	
20°	グー	0.769734	
	チョキ	0.562148	
	パー	0.467807	パー
30°	グー	0.428923	
	チョキ	0.478061	
	パー	0.221109	パー
40°	グー	0.248176	グー
	チョキ	0.587304	
	パー	0.53371	
50°	グー	0.182225	グー
	チョキ	0.489981	
	パー	0.43955	
60°	グー	0.403991	
	チョキ	0.452142	
	パー	0.309342	パー
70°	グー	1.548926	グー
	チョキ	1.60981	
	パー	1.792719	
80°	グー	1.277049	グー
	チョキ	1.415893	
	パー	1.599416	
90°	グー	0.350539	グー
	チョキ	0.548254	
	パー	0.704238	

表 4.4.1 8×8 画素において正答率が最も高かった N=28

各角度に対するユークリッド距離と回答カテゴリ(左：グー、中央：チョキ、右：パー)

角度	ラベル	距離	答え	角度	ラベル	距離	答え	角度	ラベル	距離	答え
0°	グー	0	グー	0°	グー	0.732991		0°	グー	0.193589	
	チョキ	0.732991			チョキ	0	チョキ		チョキ	0.703964	
	パー	0.193589			パー	0.703964			パー	0	パー
10°	グー	0.080366	グー	10°	グー	0.321029	グー	10°	グー	0.876581	
	チョキ	0.734524			チョキ	0.793158			チョキ	0.39585	チョキ
	パー	0.260982			パー	0.383729			パー	0.901808	
20°	グー	0.304097	グー	20°	グー	0.417833		20°	グー	0.348812	グー
	チョキ	0.489737			チョキ	0.38908			チョキ	0.964781	
	パー	0.357231			パー	0.367294	パー		パー	0.367434	
30°	グー	0.630116		30°	グー	1.087715		30°	グー	0.184657	グー
	チョキ	0.32707	チョキ		チョキ	0.617388	チョキ		チョキ	0.762431	
	パー	0.661947			パー	1.115745			パー	0.240899	
40°	グー	0.631286		40°				40°	グー	0.464104	グー
	チョキ	0.319451	チョキ						チョキ	0.563733	
	パー	0.651558							パー	0.536828	
50°	グー	0.900502		50°	グー	0.483114	グー	50°	グー	0.375125	グー
	チョキ	0.37154	チョキ		チョキ	0.974047			チョキ	0.503064	
	パー	0.921419			パー	0.505152			パー	0.414015	
60°	グー	0.859845		60°	グー	0.761461		60°	グー	0.225172	グー
	チョキ	0.350675	チョキ		チョキ	0.527226	チョキ		チョキ	0.7814	
	パー	0.872311			パー	0.807853			パー	0.345669	
70°	グー	0.35815		70°	グー	0.460804	グー	70°	グー	1.779717	
	チョキ	0.586633			チョキ	0.499923			チョキ	1.382791	チョキ
	パー	0.323699	パー		パー	0.4704			パー	1.823374	
80°	グー	0.372978		80°	グー	0.287336	グー	80°	グー	1.626481	
	チョキ	0.594219			チョキ	0.913019			チョキ	1.173742	チョキ
	パー	0.332265	パー		パー	0.36512			パー	1.683391	
90°	グー	0.222751		90°	グー	0.842074		90°	グー	0.960996	
	チョキ	0.848621			チョキ	0.527959	チョキ		チョキ	0.457184	チョキ
	パー	0.220789	パー		パー	0.822159			パー	1.001545	

表 4.4.2 8×8 画素において正答率が最も低かった N=36

各角度に対するユークリッド距離と回答カテゴリ(左：グー、中央：チョキ、右：パー)

まず表 4.3.1 の考察を行う。「グー」の誤結果として「パー」が回答として出ているが、これは比較的「パー」の形状が丸みのあるものになっているため、「パー」と誤判断しているのではないかと考えられる。逆に「パー」は全て正解しているが、20° 以上の角度がついた場合、ユークリッド距離の差が小さく、安定した結果とは言い難い。また、「チョキ」の判別は全体的に上手くいっていない。

次に 16×16 画素の結果で最も正答率の低かった N=38 の時(表 4.3.2)の考察を行う。「グー」と「パー」間における判別が上手くいっていない。また、「チョキ」を含めて回転角度が小さいものは比較的ユークリッド距離差が付くものの、全体的に結果の差が小さく曖昧である。

8×8 画素の結果を考察すると、最も高かった N=28 にしても最も低かった N=36 にしても全体的結果が安定していない。8×8 の「チョキ」の正答率が若干低くなるものの、図 4.17 のように 8×8 における「チョキ」角度 40° における輪郭線の一部(指先部分)が離れ小島のようになってしまいエラーが発生したため、40° に関しては確実に誤回答している。



図 4.17 8×8 、角度 40° 「チョキ」

以上の結果から、Z 型フーリエ記述子は低周波数帯($k = 1 \sim k = 6$)における画像の再生が十分でない可能性が考えられ、そのため、画像の特徴と成る様な指などの部分が特徴量として記述子に反映されていないものと考えられる。

4.3.2 P 型フーリエ記述子パターン認識を用いたシミュレーション

まず、第 2 章 2.4.1 (iii) で説明した偏角関数表現である P 型フーリエ記述子を用いたパターン認識シミュレーションを行う。P 型フーリエ記述子で使用する周期関数は、

$$\begin{aligned} w_i &= \exp(j\theta_i) = \cos\theta_i + j\sin\theta_i \\ &= \frac{x_{i+1} - x_i}{\delta} + j \frac{y_{i+1} - y_i}{\delta} = \frac{z_{i+1} - z_i}{\delta} \end{aligned} \quad (2.9)$$

である。シミュレーションでは Z 型フーリエ記述子の際と同様に、 $N=18 \sim 40$ の 12 種類の点数で代表点抽出を行い、評価は低周波成分の $k = 1 \sim k = 6$ までのフーリエ記述子を用いた。

例として、図 4.18 は 16×16 画素の「グー」「チョキ」「パー」 0° 基本画像のフーリエ係数である。図 4.18 の上のグラフは出力された係数の絶対値である。

また図 4.19 は 8×8 画素の「グー」「チョキ」「パー」 0° 基本画像のフーリエ係数である。図 4.18 と同様に、上のグラフは実数部、下のグラフは虚数部を示す。

次に認識シミュレーションを行うため、モデルデータの特徴ベクトルと認識対象画像の特徴ベクトルとの評価を行う。評価には、第 2 章における 2.4.4 の評価式を用い、モデル特徴ベクトルと認識画像の特徴ベクトルにおける各係数のユークリッド距離が最も小さいものを答えとして出力する。本項の表 4.5、表 4.6 で示すのは、その結果の正答率である。

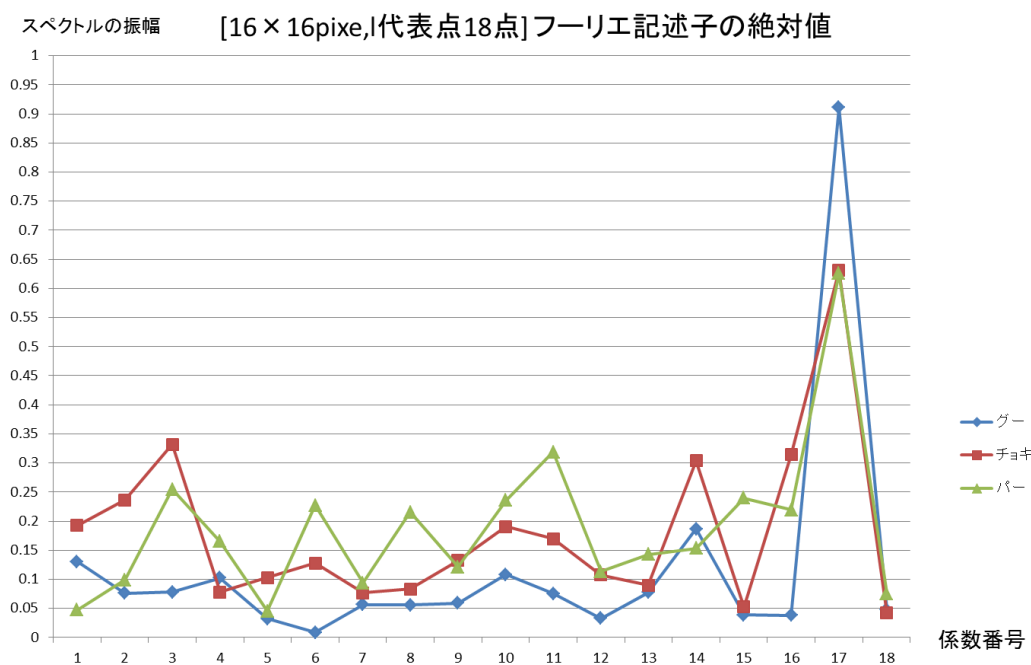


図 4.18 フーリエ記述子の振幅(16×16、代表点 18 点)

スペクトルの振幅 [8×8pixe,l代表点18点] フーリエ記述子の絶対値

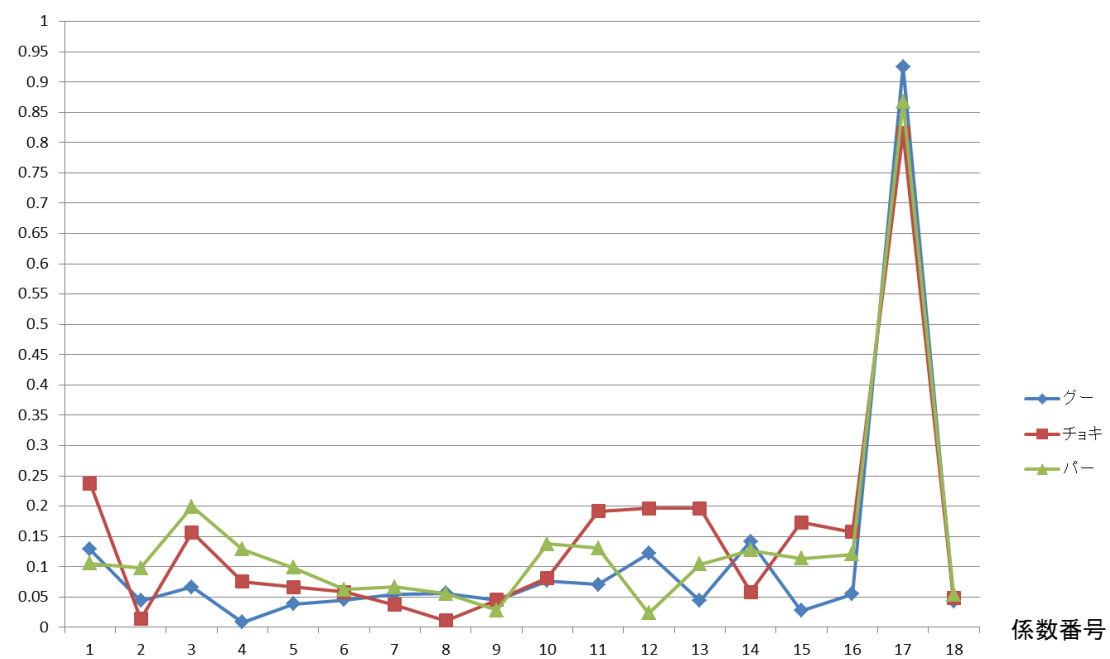


図 4.19 フーリエ記述子の振幅(8×8、代表点 18 点)

ゲー				チョキ				パー			
k	正解	不正解	正答率	k	正解	不正解	正答率	k	正解	不正解	正答率
18	10	0	100	18	10	0	100	18	10	0	100
20	10	0	100	20	10	0	100	20	8	2	80
22	10	0	100	22	10	0	100	22	10	0	100
24	10	0	100	24	10	0	100	24	10	0	100
26	10	0	100	26	10	0	100	26	10	0	100
28	10	0	100	28	10	0	100	28	10	0	100
30	10	0	100	30	10	0	100	30	10	0	100
32	10	0	100	32	10	0	100	32	10	0	100
34	10	0	100	34	10	0	100	34	10	0	100
36	10	0	100	36	10	0	100	36	10	0	100
38	10	0	100	38	10	0	100	38	10	0	100
40	10	0	100	40	10	0	100	40	10	0	100

N	総合正答率
18	100.00
20	93.33
22	100.00
24	100.00
26	100.00
28	100.00
30	100.00
32	100.00
34	100.00
36	100.00
38	100.00
40	100.00
平均	99.44

表 4.5 P 型フーリエ記述子正答表：16×16 画素における代表点 N ごとの正当結果

ゲー				チョキ				パー			
k	正解	不正解	正答率	k	正解	不正解	正答率	k	正解	不正解	正答率
18	10	0	100	18	7	3	70	18	10	0	100
20	10	0	100	20	7	3	70	20	10	0	100
22	10	0	100	22	7	3	70	22	10	0	100
24	10	0	100	24	6	4	60	24	10	0	100
26	10	0	100	26	8	2	80	26	10	0	100
28	10	0	100	28	7	3	70	28	10	0	100
30	10	0	100	30	9	1	90	30	10	0	100
32	10	0	100	32	6	4	60	32	10	0	100
34	10	0	100	34	9	1	90	34	10	0	100
36	10	0	100	36	8	2	80	36	10	0	100
38	10	0	100	38	7	3	70	38	10	0	100
40	10	0	100	40	9	1	90	40	10	0	100

N	総合正答率
18	90.00
20	90.00
22	90.00
24	86.67
26	93.33
28	90.00
30	96.67
32	86.67
34	96.67
36	93.33
38	90.00
40	96.67
平均	91.67

表 4.6 P 型フーリエ記述子正答表：8×8 画素における代表点 N ごとの正当結果

表 4.5、表 4.6 から判るように、正答率に関して非常に良い結果が出たと思われる。

形状情報が比較的多く保存され、特徴量が大きいと思われる 16×16 画素の正答率は非常に高く、全体でほぼ 100%近い正答率という結果になった。

また、8×8 においても最適な N においては最高正答率 96%、全体の平均正答率も 91% となった。8×8 の 40° 条件における「チョキ」のエラーを考慮した場合、N=30 または N=40 に関しては 100%の正答率になる可能性も十分に考えられる。

表 4.7.1 と表 4.7.2、表 4.8.1 と表 4.8.2 で、最も正答率の高い N と最も正答率の低い N に関する各ジェスチャーのユークリッド距離、回答カテゴリを示す。

角度	ラベル	距離	答え
0°	グー	0	グー
	チョキ	0.304955	
	パー	0.241032	
10°	グー	0.045954	グー
	チョキ	0.30476	
	パー	0.250604	
20°	グー	0.106033	グー
	チョキ	0.304477	
	パー	0.25654	
30°	グー	0.121346	グー
	チョキ	0.31057	
	パー	0.262109	
40°	グー	0.135099	グー
	チョキ	0.331332	
	パー	0.310146	
50°	グー	0.126937	グー
	チョキ	0.296963	
	パー	0.270006	
60°	グー	0.113886	グー
	チョキ	0.313588	
	パー	0.279884	
70°	グー	0.114375	グー
	チョキ	0.319332	
	パー	0.276938	
80°	グー	0.096102	グー
	チョキ	0.333408	
	パー	0.253506	
90°	グー	0.091871	グー
	チョキ	0.332458	
	パー	0.260374	

角度	ラベル	距離	答え
0°	グー	0.304955	
	チョキ	0	チョキ
	パー	0.254008	
10°	グー	0.308897	
	チョキ	0.047763	チョキ
	パー	0.251568	
20°	グー	0.291333	
	チョキ	0.074415	チョキ
	パー	0.273919	
30°	グー	0.288082	
	チョキ	0.087916	チョキ
	パー	0.189365	
40°	グー	0.298676	
	チョキ	0.109554	チョキ
	パー	0.189631	
50°	グー	0.342901	
	チョキ	0.118081	チョキ
	パー	0.214908	
60°	グー	0.332677	
	チョキ	0.105288	チョキ
	パー	0.232934	
70°	グー	0.305196	
	チョキ	0.089362	チョキ
	パー	0.210621	
80°	グー	0.322882	
	チョキ	0.103701	チョキ
	パー	0.270569	
90°	グー	0.339524	
	チョキ	0.122208	チョキ
	パー	0.257378	

角度	ラベル	距離	答え
0°	グー	0.241032	
	チョキ	0.254008	
	パー	0	パー
10°	グー	0.299363	
	チョキ	0.278526	
	パー	0.129892	パー
20°	グー	0.241641	
	チョキ	0.28793	
	パー	0.090702	パー
30°	グー	0.230202	
	チョキ	0.284963	
	パー	0.136723	パー
40°	グー	0.315212	
	チョキ	0.324739	
	パー	0.106549	パー
50°	グー	0.318673	
	チョキ	0.284206	
	パー	0.148878	パー
60°	グー	0.229601	
	チョキ	0.302807	
	パー	0.087363	パー
70°	グー	0.36999	
	チョキ	0.284866	
	パー	0.234323	パー
80°	グー	0.353429	
	チョキ	0.309356	
	パー	0.186285	パー
90°	グー	0.373086	
	チョキ	0.308597	
	パー	0.183023	パー

表 4.7.1 16×16 画素において正答率が最も高かった N=18

各角度にたいするユークリッド距離と回答カテゴリ(左：グー、中央：チョキ、右：パー)

角度	ラベル	距離	答え
0°	グー	0	グー
	チョキ	0.295811	
	パー	0.325492	
10°	グー	0.0403	グー
	チョキ	0.320201	
	パー	0.356004	
20°	グー	0.074977	グー
	チョキ	0.313149	
	パー	0.345317	
30°	グー	0.087812	グー
	チョキ	0.327442	
	パー	0.3649	
40°	グー	0.12703	グー
	チョキ	0.340509	
	パー	0.404322	
50°	グー	0.109291	グー
	チョキ	0.31269	
	パー	0.384767	
60°	グー	0.113577	グー
	チョキ	0.320651	
	パー	0.404141	
70°	グー	0.100194	グー
	チョキ	0.317634	
	パー	0.390971	
80°	グー	0.105763	グー
	チョキ	0.335745	
	パー	0.34556	
90°	グー	0.088933	グー
	チョキ	0.336073	
	パー	0.352292	

角度	ラベル	距離	答え
0°	グー	0.295811	
	チョキ	0	チョキ
	パー	0.339085	
10°	グー	0.258537	
	チョキ	0.091875	チョキ
	パー	0.366205	
20°	グー	0.25088	
	チョキ	0.126624	チョキ
	パー	0.337379	
30°	グー	0.283228	
	チョキ	0.097118	チョキ
	パー	0.318645	
40°	グー	0.295291	
	チョキ	0.075717	チョキ
	パー	0.31014	
50°	グー	0.301943	
	チョキ	0.083691	チョキ
	パー	0.300401	
60°	グー	0.298392	
	チョキ	0.06364	チョキ
	パー	0.298439	
70°	グー	0.292399	
	チョキ	0.053198	チョキ
	パー	0.297219	
80°	グー	0.272777	
	チョキ	0.051261	チョキ
	パー	0.312182	
90°	グー	0.312359	
	チョキ	0.090702	チョキ
	パー	0.339355	

角度	ラベル	距離	答え
0°	グー	0.325492	
	チョキ	0.339085	
	パー	0	パー
10°	グー	0.348674	
	チョキ	0.371423	
	パー	0.062356	パー
20°	グー	0.224009	
	チョキ	0.372494	
	パー	0.213353	パー
30°	グー	0.197615	グー
	チョキ	0.315074	
	パー	0.205189	
40°	グー	0.217005	グー
	チョキ	0.372353	
	パー	0.236269	
50°	グー	0.215418	
	チョキ	0.325195	
	パー	0.171195	パー
60°	グー	0.239413	
	チョキ	0.32275	
	パー	0.160715	パー
70°	グー	0.388222	
	チョキ	0.366458	
	パー	0.107984	パー
80°	グー	0.348665	
	チョキ	0.405431	
	パー	0.150109	パー
90°	グー	0.355157	
	チョキ	0.375469	
	パー	0.123858	パー

表 4.7.2 16×16 画素において正答率が最も低かった N=20

各角度にたいするユークリッド距離と回答カテゴリ(左: グー、中央: チョキ、右: パー)

角度	ラベル	距離	答え
0°	グー	0	グー
	チョキ	0.242632	
	パー	0.173501	
10°	グー	0.023213	グー
	チョキ	0.254231	
	パー	0.184911	
20°	グー	0.020937	グー
	チョキ	0.24021	
	パー	0.173967	
30°	グー	0.03607	グー
	チョキ	0.223324	
	パー	0.149518	
40°	グー	0.038797	グー
	チョキ	0.228195	
	パー	0.153874	
50°	グー	0.022022	グー
	チョキ	0.24152	
	パー	0.169576	
60°	グー	0.015578	グー
	チョキ	0.247043	
	パー	0.171245	
70°	グー	0.03419	グー
	チョキ	0.251711	
	パー	0.173242	
80°	グー	0.054786	グー
	チョキ	0.25388	
	パー	0.160224	
90°	グー	0.063989	グー
	チョキ	0.275784	
	パー	0.187714	

角度	ラベル	距離	答え
0°	グー	0.242632	
	チョキ	0	チョキ
	パー	0.146516	
10°	グー	0.301611	
	チョキ	0.135259	チョキ
	パー	0.156517	
20°	グー	0.169899	
	チョキ	0.144553	チョキ
	パー	0.160125	
30°	グー	0.195924	
	チョキ	0.112167	チョキ
	パー	0.146424	
40°			
50°	グー	0.203148	
	チョキ	0.078724	チョキ
	パー	0.124241	
60°	グー	0.19352	
	チョキ	0.083844	チョキ
	パー	0.16242	
70°	グー	0.228265	
	チョキ	0.127716	チョキ
	パー	0.217341	
80°	グー	0.178182	
	チョキ	0.137873	チョキ
	パー	0.181315	
90°	グー	0.252465	
	チョキ	0.047966	チョキ
	パー	0.134765	

角度	ラベル	距離	答え
0°	グー	0.173501	
	チョキ	0.146516	
	パー	0	パー
10°	グー	0.16046	
	チョキ	0.124191	
	パー	0.08583	パー
20°	グー	0.182381	
	チョキ	0.17074	
	パー	0.066397	パー
30°	グー	0.238644	
	チョキ	0.165053	
	パー	0.097827	パー
40°	グー	0.21269	
	チョキ	0.196064	
	パー	0.095371	パー
50°	グー	0.175305	
	チョキ	0.150382	
	パー	0.044314	パー
60°	グー	0.199915	
	チョキ	0.182326	
	パー	0.085502	パー
70°	グー	0.256641	
	チョキ	0.192081	
	パー	0.133386	パー
80°	グー	0.180588	
	チョキ	0.220022	
	パー	0.084985	パー
90°	グー	0.169888	
	チョキ	0.20327	
	パー	0.079117	パー

表 4.8.1 8×8 画素において正答率が最も高かった N=30

各角度にたいするユークリッド距離と回答カテゴリ(左：グー、中央：チョキ、右：パー)

角度	ラベル	距離	答え
0°	グー	0	グー
	チョキ	0.200017	
	パー	0.165258	
10°	グー	0.030829	グー
	チョキ	0.212038	
	パー	0.184376	
20°	グー	0.055742	グー
	チョキ	0.211489	
	パー	0.182766	
30°	グー	0.06967	グー
	チョキ	0.213205	
	パー	0.167732	
40°	グー	0.059155	グー
	チョキ	0.216318	
	パー	0.172703	
50°	グー	0.083681	グー
	チョキ	0.231241	
	パー	0.182566	
60°	グー	0.093295	グー
	チョキ	0.246275	
	パー	0.182915	
70°	グー	0.091615	グー
	チョキ	0.244241	
	パー	0.167573	
80°	グー	0.104639	グー
	チョキ	0.24111	
	パー	0.152614	
90°	グー	0.108938	グー
	チョキ	0.281659	
	パー	0.194797	

角度	ラベル	距離	答え
0°	グー	0.200017	
	チョキ	0	チョキ
	パー	0.139849	
10°	グー	0.261738	
	チョキ	0.138017	
	パー	0.128149	パー
20°	グー	0.142501	グー
	チョキ	0.143464	
	パー	0.163813	
30°	グー	0.180608	
	チョキ	0.101722	チョキ
	パー	0.139112	
40°			
50°	グー	0.157575	
	チョキ	0.082865	チョキ
	パー	0.102414	
60°	グー	0.159734	
	チョキ	0.067167	チョキ
	パー	0.119911	
70°	グー	0.191056	
	チョキ	0.090735	チョキ
	パー	0.172064	
80°	グー	0.161222	
	チョキ	0.142713	チョキ
	パー	0.173757	
90°	グー	0.218839	
	チョキ	0.111461	
	パー	0.076365	グー

角度	ラベル	距離	答え
0°	グー	0.165258	
	チョキ	0.139849	
	パー	0	パー
10°	グー	0.152281	
	チョキ	0.123996	
	パー	0.098375	パー
20°	グー	0.177357	
	チョキ	0.159184	
	パー	0.037294	パー
30°	グー	0.199152	
	チョキ	0.180096	
	パー	0.075829	パー
40°	グー	0.179468	
	チョキ	0.207597	
	パー	0.092319	パー
50°	グー	0.149795	
	チョキ	0.171197	
	パー	0.063455	パー
60°	グー	0.164193	
	チョキ	0.189676	
	パー	0.089479	パー
70°	グー	0.238085	
	チョキ	0.209626	
	パー	0.110067	パー
80°	グー	0.180218	
	チョキ	0.224242	
	パー	0.091629	パー
90°	グー	0.164954	
	チョキ	0.207528	
	パー	0.077008	パー

表 4.8.2 8×8 画素において正答率が最も低かった N=24

各角度にたいするユークリッド距離と回答カテゴリ(左：グー、中央：チョキ、右：パー)

表 4.7.1、表 4.7.2、表 4.8.1、表 4.8.2 の全体として評価できると思われる点は、まず、答えとした最短距離の値とそれ以外の値と間で比較的值に差があるということである。これは、各要素のスペクトルの振幅に特徴づける差が発生していることが予想される。

また、回転による画像影響を考慮した上でも影響が少なく、回転に対する耐性があると言える。

特筆して言えるのは、8×8 画素における「チョキ」の正答率である。他のジェスチャーに比べれば正答率こそ劣るものの、図 4.10 で解るように、基本となる 0° 画像と他の角度のものとは若干異なるものとなっている。これは、低解像度画像における問題の一つである、回転処理における近傍画素との輝度値変化の影響が大きい、ということに起因する。正面画像では中指と人差し指は分離されているが、角度が変わることによって、近傍画素との間で平滑化が行われて「二股に分離している」という情報が失われてしまっている。視覚的に判別する際に、最も判別し易いと思われる指情報が欠損して

しまった状態で、にも拘らず判別出来ている理由として次のことが考えられる。指情報では無く、大まかな形状の差異で判別しているのではないか、ということだ。「グー」はどの回転角度であっても比較的丸に近い形状を保っている。「パー」は比較的眞四角に近い形を取っており、また、親指の輪郭部で大きな変化がある。そして「チョキ」は全体的に長方形に近い形を取っているように見える。今回のシミュレーションにおける評価は $k=1\sim6$ と比較的低い周波数成分のみを用いている。画像は高周波成分を伴うにつれて角のついたメリハリのある画像として再生されていく(図 4.9)わけだが、 8×8 の場合、回転による近傍画素との影響が非常に大きいことから、同じ対象のジェスチャーであっても回転によって高周波成分のスペクトルが大きく変化する可能性がある。つまり、「チョキ」を例にして言えば、指部分の情報が再生される回数に達していなかったため、逆に判別が上手くいったのではないか、ということだ。指の情報が高周波成分によってその特徴を上手く再生されてしまった場合、 0° と他の角度とでは形状に差があると言わざるを得ない。その場合、結果としてユークリッド距離が離れてしまうということが考えられる。とはいえ、ここまでの結果から低周波成分の差だけでもジェスチャーを十分に判別できる、とも言えるだろう。

16×16 画素では回転しても輪郭情報に差は小さいので、評価する次数の値を大きくしても十分な結果が出るであろうことが予想される。

総じて、P 型フーリエ記述子は「グー」「チョキ」「パー」の回転画像に対して、比較的低い次数のみを評価することで、我々が意図した結果を回答してくれるのではないかと考えられる。

DCT 係数のシミュレーション結果と異なり、正面 0° のモデルだけ用意しておけば、回転などの変化に対応できる。また、アルゴリズムの性質上、拡大や平行移動にも対応できると考えられるが、低解像度画像であるが故に画像の一部が画像サイズ外に漏れてしまった場合や、非常に対象が小さ過ぎてうまく画像抽出できない場合などでは、上手くいかない可能性がある。

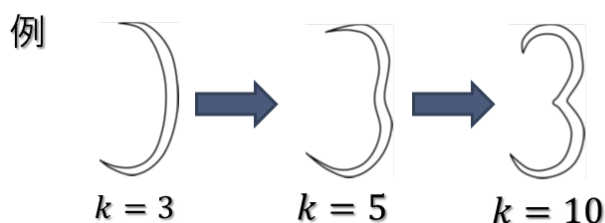


図 4.9 「3」を例とした低周波成分と高周波成分の関係

第 5 章

結論と展望

5.1 研究結果のまとめ

本研究では、まず熱赤外センサを用いた非接触 UI のシステム開発を提案した。その上で、現状の熱赤外センサの状況を鑑みて低解像度空間におけるパターン認識の必要性を確認し、実際に熱赤外低解像度画像におけるジェスチャー認識シミュレーションの実現性とアルゴリズムなどのシステムにおける基礎となる部分の確認を目的とした。

簡便な処理を目指し、データベースに登録するモデル数をできるだけ削減し、なおかつ安定した判別ができるよう、画像からのジェスチャー抽出処理方法及びパターン認識アルゴリズムの選定に関する考察を行った。

その結果、DCT 係数を用いたパターン認識手法では、回転の動作などの画像の撮影条件の変化に弱く安定性に欠けるという結果が得られた。次に行った Z 型フーリエ記述子を用いたパターン認識手法におけるシミュレーション判別では、低い次数において望むような収束結果が得られず、アルゴリズムの見直しおよびデータ入力法・抽出法の見直す必要があると考えられる。

最後に行った、P 型記述子を用いたパターン認識手法は今回用いた解像度(16×16 画素、8×8 画素)では比較的高い正答率を得ることができ、低解像度空間においても回転に対して安定した判別が行えた。8×8 画素におけるジェスチャー「チョキ」において正答率が若干下がってしまったが、この点に関しては低解像度空間における輪郭抽出法の改善によって、正答率を向上することが可能であると思われる。全体的に低い次数のみで判別することができる。これはシステムの高速化に繋り、好ましい結果であったと考えられる。

また、ある時間区間内における温度情報の著しい変化なども非接触ユーザ・インターフェースの信号として使用することも検討していくべきだと考える。例えば、ある設定したクロックで手を翳して温度を明滅させることで、それをトリガーとする。この場合、細かい情報を考慮する必要もなく、パターン信号とすることが出来る。今後はそういった手法も合わせながらシステム構築を考えていきたい。

これらのことから、まだ改良の余地はあるものの熱赤外の低解像度画像においてもジェスチャー認識が可能であるということを確認しつつ、今後もこの低解像度熱赤外線を用いたパターン認識のシステムが確立される可能性は十分あると考える。

5.2 今後の展望

回転に対して安定性が高いと P 型フーリエ記述子を用いたパターン認識手法において、「手のひらの傾き」という条件に対して、どれほどの許容性があるかの確認をしておくべきだと考える。また、特異なデータにどのように対処するか、などアルゴリズム段階での検討の余地がまだあるので、その点を今後詰めていく必要があると考える。

また、本研究の最終目的はセンサから処理を含めた「システム開発」であることから、どのような機器を用いて、どのように画像を取得するか、等の包括的な見方で研究を進めていく必要がある。今後、実際に安価に試作できるような熱赤外線センサを作成し、その上でアルゴリズム等を発展させていくべきだろうと考える。

謝辞

本研究を進めるにあたり、指導を頂いた伊藤直史先生、稲村實先生に深く感謝致します。研究テーマの模索に当たり大変有益な助言を頂き、このテーマで研究を進めることが出来、今では大変安堵しております。また、こうして研究を進められたのも研究室の皆様のおかげだったと思っております。この場を借りて心より感謝致します。

参考文献

- 1) 酒井幸市. 画像処理とパターン認識入門. 森北出版, 12 2006.
- 2) 大類重範. デジタル信号処理. 日本理工出版会, 12 2001.
- 3) 斎藤恒雄. 画像処理アルゴリズム. 近代科学者, 3 1993.
- 4) Richard O.Duda, Peter E. Hart, David G. Stork. Pattern Classification Second Edition . 新技術コミュニケーションズ, 7 2001.
- 5) 高木幹雄, 下田陽久. 画像解析ハンドブック. 東京大学出版会, 1 1991.
- 6) 安居院猛, 長尾智晴. C 言語による画像処理入門. 昭晃堂, 11 2000.
- 7) 林晴比古, 新訂 C 言語入門ビギナー編. ソフトバンククリエイティブ, 12 2003.
- 8) 林晴比古, 改訂 新 C 言語入門応用編. ソフトバンクパブリッシング, 9 1999.
- 9) (株)アंक, C の絵本. 翔泳社, 3 2002.
- 10) 岡谷 貴之, コンピュータビジョン最先端ガイド 3. アドコム・メディア, 12 2010.
- 11) 阿部 拓也, 超低解像ナンバープレート画像の文字認識. 2010年度 稲村實研究室卒業論文